LASER DIFFRACTION PARTICLE SIZING: SAMPLING AND INVERSION

by
James B. Riley
B.S. Physics, Yale University, 1981

SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
and the
WOODS HOLE OCEANOGRAPHIC INSTITUTION
August 13, 1987

Signature of Author_____.

Joint Program in Oceanographic Engineering
Massachusetts Institute of Technology
Woods Hole Oceanographic Institution

Certified by_____.

Y. C. Agrawal
Thesis Supervisor

Accepted by_____.

George V.Frisk
Chairman, Joint Committee for Oceanographic Engineering,
Massachusetts Institute of Technology
Woods Hole Oceanographic Institution

Abstract

The inverse problem of obtaining particle size distributions from observations of the angular distribution of near forward scattered light is reexamined. Asymptotic analysis of the forward problem reveals the information content of the observations, and the sources of non-uniqueness and instability in inverting them. A sampling criterion, such that the observations uniquely specify the size distribution is derived, in terms of the largest particle size, and an angle above which the intensity is indistinguishable from an asymptote. The instability of inverting unevenly spaced data is compared to that of super-resolving Fourier spectra. Resolution is shown to be inversely proportional to the angular range of observations.

The problem is rephrased so that the size weighted number density is sought from the intensity weighted by the scattering angle cubed. Algorithms which impose positivity and bounds on paritcle size improve the stability of inversions. The forward problem can be represented by an over-determined matrix equation by choosing a large integration increment in size dependent on the frequency content of the angular intensity, further improving stability.

Experimental data obtained using a linear CCD array illustrates the

theory, with standard polystyrene spheres as scatterers. The scattering from single and tri-modal distributions is successfully inverted.

Acknowledgements

By far, my thesis advisor Y. C. Agrawal has assisted me most, in this work and in my graduate education; for this, I thank him. Few students learn as much from their advisors as I have. Yogi has always been accessible and interested in my progress. His creativity provided the inspiration for this thesis. Of the two of us, Yogi first noticed that the function $\theta^3 I(\theta)$ of the scattered intensity from particles, is, in the diffraction approximation, asymptotically periodic; this underlies the entire thesis. His ideas made the experimental instrumentation work.

I am especially grateful to the nucleus of graduate students studying inverse theory at Woods Hole. I have profited immensely from discussions with them; thank you, C. S. Chiu, J. H. Miller, and S. D. Rajan.

I thank the thesis committee, of J. F. Lynch, A. V. Oppenheim and A. J. Williams III for their valuable comments and suggestions, and G. V. Frisk for chairing the thesis defense.

A number of individuals assisted considerably in the experimental work. Chris Belting helped me with the electronics, especially with keeping the CCD photodiode array working. M. C. Woodward performed much of the machining, and provided valuable advice on mechanical design and

machining. P. G. Brewer, P. L. Sachs, and A. L. Bradshaw provided the analog to digital converter. A. J. Williams III provided a wealth of useful information.

Finally, I thank my family for support and encouragement.

During the work represented here, I was supported by a NASA Technology Transfer Traineeship grant (NGT-014-800, Supplement 5), and by the Joint Program in Oceanographic Engineering between the Woods Hole Oceanographic Institution and the Massachusetts Institute of Technology. The experimental work was funded by the Coastal Research Laboratory of the Woods Hole Oceanographic Institution, through the generosity of the Mellon Foundation.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In many physical processes in the ocean, the size distribution of suspended particles has central importance. In sediment transport, it is an essential term in the equations of momentum and mass continuity. For light scattering and distribution, the size distribution largely determines the difference between pure and ocean water. The same is true for acoustic scattering.

Despite its importance, measurements of particulate size distribution in the ocean are scarce and incomplete. Most measurements have been made with Coulter counters. The data is poor in resolution; there is virtually no data for particles smaller than one micron. There have been no measurements yet in which particulate size distribution is measured temporally, with enough resolution in time and space for the study of the *dynamics* of sediment transport.

In this work, I seek to rectify the shortage of particle size distribution measurements. I consider the design of an instrument that measures particle size distribution, based on a method called *optical diffraction particle sizing*. The method consists of inverting measurements of the angular dis-

tribution of near-forward scattering from a collimated beam. It has the advantage of being able to measure size distribution non-intrusively and nearly instantaneously, allowing time series of size distribution to be obtained.

Optical diffraction particle sizing is based on an *inverse problem.* It is based on inverting an integral of the form

$$I(\theta) = \int_0^\infty n(x) I(\theta, x)\, dx, \tag{1.1}$$

where $I(\theta)$ represents the scattering from a collimated beam into an angle θ, from an ensemble of particles with number density $n(x)$. Here x is the particle radius, non-dimensionalized by wave-number; that is, $x = kr$, where r is the particle radius, and k is the optical wave-number. The function $I(\theta, x)$ represents the scattering from a single particle of radius x into an angle θ.

In this thesis, I focus on the issues of instability and non-uniqueness in the inversion of experimental measurements of scattering. These include not just the fundamental, mathematical ill-posedness of inverting equation (1.1) but the numerical problems of inverting discrete samples of $I(\theta)$ from a limited range of angles.

My approach is to perform an asymptotic analysis of equation (1.1), using diffraction theory to model the kernel $I(\theta, x)$. This shows that the angular distribution of scattering, $I(\theta)$, and the particle number density, $n(x)$, are approximately related by a Fourier-sine transform. This results in new understanding of the inverse problem, in terms of familiar concepts of Fourier transforms, such as the Nyquist sampling criterion (or Shan-

non sampling theorem), the Heisenberg uncertainty principle, and super-resolution and band-width extrapolation of band-limited images. I apply this understanding in deriving sampling criteria, in comparing algorithms, and in applying *a priori* information for inverting $I(\theta)$. For this, I use experimental, rather than synthetic data, from measurements of scattering from polystyrene spheres with a prototype instrument of my design. Finally, I apply the theory to measurements of scattering from natural suspensions, to provide data for the design of a particle sizing instrument for oceanographic work.

Though the theory herein has been developed with physical processes in the ocean in mind, it has broader applicability. It has value for sizing aerosol sprays in combustion processes, for monitoring environmental pollutants, and for medical applications, for example.

Related Work Malvern Instruments, Ltd. of Malvern, England markets an optical diffraction particle sizing instrument [4,5,6]. Despite the availability of a commercial instrument, there is considerable interest in the academic community in designing and using diffraction particle sizers, and in improving the method.

Some of the academic interest results from the failure of the Malvern instrument to produce results which are comparable to those of other instruments for certain types of particles. McCave, for example, has compared measurements from the Malvern instrument with those of a Coulter counter on clay from the Nova Scotian rise [7]. The clay was collected in the High Energy Benthic Boundary Layer Experiment (HEBBLE). The re-

sults of the two methods do not consistently agree, and for the Malvern instrument, do not agree for focusing lenses of differing focal lengths. The function of the focusing lenses will be discussed in section 1.5. Differing focal lengths correspond to differing samplings of scattering in angle, both in sampling interval and range. The significance of sampling rate and range is a major focus of this work, and may explain some of the shortcomings of the Malvern instrument.

A number of investigators have studied an analytical inverse to (1.1), derived by Chin, Sleipcevich, and Tribus [8,9]. This is based on diffraction theory, to be discussed in chapter 2:

$$n(x) = \frac{-2\pi k^3}{x^2} \int_0^\infty J_1(x\theta) Y_1(x\theta) x\theta \, \frac{d}{d\theta} \left[\theta^3 I(\theta)\right] \, d\theta \qquad (1.2)$$

Here J_1 and Y_1 are first order Bessel functions of the first and second kind. These investigators have studied the effects of sampling over a limited range of angles, and of varying the number of sample points of $I(\theta)$, on this integral. Some of their work is based on empirical, computer simulations of the scattering from synthesized size distributions [10,11]. Other work is based on inverting size distributions of a specific form, and inverting analytically [12,13]. The disadvantages of the analytical inverse are several, and are discussed in chapter 3.

The work I present here is similar but fundamentally different, in that it applies to matrix inversions of equation (1.1), and to the evaluation of the analytical integral of Chin *et. al* as well.

Diffraction particle sizing has had limited use in the ocean. Shifrin [14] and Sokolov [15] have reported on *in situ* instruments for measurements

in the sea, but little of their work has been reported in western literature. Their work is also based on the analytical inverse.

1.1 Outline of Thesis

My interest in diffraction particle sizing stems from an interest in several physical processes in the ocean. Thus, in section 1.2 I present two examples of physical processes in the ocean where size distribution is an important parameter. Section 1.3 surveys the literature of particle distributions in the ocean, illustrating the scarcity of data and providing background for the theoretical undertaking of later sections. I compare methods other than diffraction particle sizing in section 1.4, and describe diffraction particle sizing in section 1.5.

Chapter 2 supplies background material. For this problem, this is the forward problem, the scattering of light from small particles. Three theories of scattering from particles are useful: the exact, Mie theory of scattering from spheres, diffraction theory, and Rayleigh scattering.

Chapter 3 lays the theoretical foundation of the thesis. I show that a function of the angular scattered intensity, $\theta^3 I(\theta)$, and of the size distribution, $xn(x)$, are asymptotically related by a Fourier-sine transform, which is more stable than the transform relating $I(\theta)$ and $n(x)$. By exploiting the asymptotic relation to a Fourier transform, I derive a Nyquist criterion based on the largest size in a distribution and the sampling interval in θ. By comparing the optical diffraction particle sizing problem to the Fourier-sine transform, I illustrate the instability of inverting scattering measurements

which have missing or unevenly spaced data. These results, based on the diffraction approximation, are approximately valid for the scattering via other models as well, e.g. Mie scattering from spheres.

Chapter 4 applies the conclusions of chapter 3 to the numerical problems of finding size distribution from angular scattered intensity and vice versa. For the special case of measurements with no missing data, I show how to construct a matrix that describes the scattering from an assembly of particles that is nearly non-singular, yet physically complete. This allows for efficient computation of $I(\theta)$ from $n(x)$ and vice-versa. I also compare some inversion algorithms, in light of the underlying causes of instability. Some common algorithms used in inverse problems are found to be inappropriate.

Chapter 5 begins the second part of this work, the experimental component. This section describes a prototype instrument measuring scattering at small angles, laboratory procedure, and initial processing of data.

In chapter 6 the theory is applied to the experimental measurements of scattering. I invert the scattering from distributions of polystyrene spheres, and present measurements of scattering from natural samples as well, to illustrate the sampling requirements for measuring the scattering from naturally occuring distributions.

Finally, chapter 7 summarizes the contributions of this work and suggests further research.

1.2 Importance of Size Distributions in the Ocean: Some Examples

The size distribution of suspended particles is an important parameter in the physics of many physical processes in the ocean. I present two examples: the fluid dynamics of boundary layers, and the scattering and distribution of light.

1.2.1 Hydrodynamics: Particles in Boundary Layers

Most sediment transport in the ocean occurs in turbulent bottom boundary layers. Suspended sediment modifies bottom boundary layer flows, through the displacement of water. Through this displacement, a suspension of particles induces a gradient in fluid density, which competes with mixing due to turbulence. The results of this competition are a stably stratified flow, with the heaviest particles concentrated nearest the bottom boundary.

The displacement of water requires modification of the equations of fluid motion. In the paragraphs that follow, I show this, deriving the Reynolds averaged Navier-Stokes equations for a bottom boundary layer, but do not attempt to solve the equations. Solution requires additional modelling, for turbulent closure; this is shown in the literature [16,17,18,19,20]. The modified equations of motion highlight size distribution as a relevant parameter that determines how particles affect boundary layer flows.

In the derivation, one assumes that the particle concentration is high enough to consider the volume concentration, denoted by $\nu(r)$, large enough

to be considered a continuum. The "fluid" represents both the water and the particles.

The presence of the suspension requires modification of the equation of mass continuity. Following Hunt [16], the volume concentration of particles of size r is related to the divergence of particle flux, $\mathbf{p}(r)$:

$$\frac{\partial \nu(r)}{\partial t} = -\nabla \cdot \mathbf{p}(r) \tag{1.3}$$

I denote the flux of the water as $\mathbf{q}(r)$, as distinguished from the particle flux, and the total fluid flux, $\mathbf{q} + \int \mathbf{p}(r)\, dr$. The total fluid is incompressible:

$$\nabla \cdot \left(\mathbf{q} + \int_0^\infty \mathbf{p}(r) dr \right) = 0. \tag{1.4}$$

Equations (1.3) and (1.4) are the equations of mass conservation.

As is usual in deriving Reynolds averaged Navier Stokes equations, the water velocity, denoted by the vector $\mathbf{u}_w$, and particle velocity denoted by $\mathbf{u}_p(r)$, are written in terms of steady and random parts:

$$\mathbf{u}_w = \overline{\mathbf{u}}_w + \mathbf{u}'_w \tag{1.5}$$

$$\mathbf{u}_p(r) = \overline{\mathbf{u}}_p(r) + \mathbf{u}'_p(r) \tag{1.6}$$

where the steady components have over-bars and the random components are primed. Similarly the particle volume concentration can be written in terms of random and steady parts:

$$\nu(r) = \overline{\nu}(r) + \nu'(r) \tag{1.7}$$

Next, assume the flow in a boundary layer to be horizontally invariant; only vertical gradients exist. Further, the particles have no inertia with

respect to the water, but have a velocity which differs only by a sinking velocity, $w_s(r)$. Thus, writing vectors $\mathbf{u}$ as $u\hat{\mathbf{x}} + v\hat{\mathbf{y}} + w\hat{\mathbf{z}}$, where $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$ are the unit vectors of a three dimensional coordinate system,

$$\overline{u'_w} = 0 \tag{1.8}$$

$$\overline{v}_w = 0,\ \overline{v'_w} = 0 \tag{1.9}$$

$$\overline{u}_p(r) = 0,\ \overline{u'_p(r)} = 0 \tag{1.10}$$

$$\overline{v}_p(r) = 0,\ \overline{v'_p(r)} = 0 \tag{1.11}$$

$$\overline{w}_p(r) = \overline{w}_w - w_s(r),\ \overline{w'_p}(r) = 0 \tag{1.12}$$

Thus the fluxes $\mathbf{q}$ and $\mathbf{p}(r)$ are

$$\mathbf{q} = w_w \left(1 - \int_0^\infty \nu(r)\, dr\right) \hat{\mathbf{z}} \tag{1.13}$$

$$\mathbf{p}(r) = (w_w - w_s)\nu(r)\hat{\mathbf{z}} \tag{1.14}$$

Reynolds (ensemble) averaged, with the vertical components of equations (1.5) and (1.6) substituted in, these fluxes become:

$$\langle q \rangle = \overline{w}_w \left(1 - \int_0^\infty \overline{\nu(r)}\, dr\right) - \int_0^\infty \overline{w'\nu'(r)}\, dr \tag{1.15}$$

$$\langle p(r) \rangle = (\overline{w}_w - w_s)\, \overline{\nu(r)} - \overline{w'_w \nu'(r)} \tag{1.16}$$

where $\langle q \rangle$ and $\langle p(r) \rangle$ now stand for the horizontal components of water and particle flux, respectively.

Substituting the latter two equations into, (1.3) and (1.4) yield the Reynolds averaged equations of mass conservation:

$$\frac{\partial \overline{\nu(r)}}{\partial t} = -\frac{\partial}{\partial z} \left\{ (\overline{w}_w - w_s(r))\, \overline{\nu(r)} + \overline{\nu'(r) w'_w} \right\} \tag{1.17}$$

and

$$\frac{\partial \int \overline{\nu(r)}\, dr}{\partial t} = \frac{\partial}{\partial z}\left\{\overline{w}_w \left(1 - \int \overline{\nu(r)}\, dr\right) - \int \overline{w'_w \nu'(r)}\, dr\right\} \tag{1.18}$$

These, along with the Reynolds averaged momentum equations,

$$\frac{\partial \overline{u}_w}{\partial t} - f\overline{v}_w = \frac{\partial}{\partial z}\left(-\overline{u'_w w'_w}\right) \tag{1.19}$$

and

$$\frac{\partial \overline{u}}{\partial t} - f\overline{u}_w = \frac{\partial}{\partial z}\left(-\overline{v'_w w'_w}\right) \tag{1.20}$$

form the equations of motion. Here I include the Ekman terms $f\overline{u}$ and $f\overline{v}$ for benthic boundary layers, where f is the Coriolis parameter.

Unfortunately, this set of equations is not complete; there is the usual problem of turbulent closure. Several investigators have tackled the problem of turbulent closure [21,18,16,17,19,20]. Their approach is to treat the turbulent flux $\overline{w'_w \nu'(r)}$ as a quantity that diffuses at a rate proportional to the gradient of $\overline{\nu(r)}$:

$$\overline{w'_w \nu'(r)} = -\epsilon_p \nabla \overline{\nu(r)} \tag{1.21}$$

where ϵ_p is a diffusivity, and similary for the Reynolds stresses $\overline{u'v'}$:

$$\overline{u'_w w'_w} = -\epsilon_w \frac{\partial}{\partial z}\overline{u(r)} \tag{1.22}$$

where ϵ_w is the familiar eddy diffusivity.

Although equations (1.17), (1.18), (1.19), and (1.20), are not complete, these equations indicate the parameters through which suspended particles affect fluid flow. These are $w_s(r)\overline{\nu(r)}$, the concentration weighted fall velocity, and $\overline{\nu'(r)w'_w(r)}$. Since these contain the particle size distribution

explicitly, particle size distribution is clearly important in understanding the effects of suspended sediment in bottom boundary layers.

While investigators have achieved some apparent success tackling the problem of turbulent closure, and have developed complicated models of sediment transport, there have been few, if any experimental measurements to which the models can be compared. Adams and Weatherly note that, (for Benthic Boundary Layers)

> We are unaware of any definitive BBL observations, i.e. one that has yielded simultaneous profiles of current speed and suspended sediment concentration, to which we can compare the output of our model. [21]

The lack of dynamic measurements of size distributions motivates, in part, this thesis.

1.2.2 Effects on Light Scattering

Suspended particles dominate light scattering, absorption and attenuation in the ocean. Optical properties of the ocean are affected by particle concentration, size distribution, index of refraction, and shape distribution. Less important are the Rayleigh scattering of water molecules, and fluctuations of index of refraction.

Practitioners of optical oceanography characterize the optical properties of ocean waters most often by the following:

- the absorption coefficient, a: the energy fraction of a collimated beam that is absorbed per unit length;

- the scattering coefficient, b; the energy fraction of a collimated beam that is scattered per unit length;

- the attenuation coefficient, α; the sum of a and b, the total attenuation per unit length;[1]

- the volume scattering function, $\beta(\theta)$,

$$\beta(\theta) = \frac{dI(\theta)}{d\theta} dV \tag{1.23}$$

  the scattered radiant intensity in a direction θ per unit scattering volume dV for unit incident irradiance.

The scattering coefficient b is related to the volume scattering function through the integral

$$b = 2\pi \int_0^\pi \beta(\theta) \sin\theta \, d\theta. \tag{1.24}$$

The volume scattering function depends explicitly on particle size distribution through the integral

$$\beta(\theta) = \int_0^\infty n(x)\sigma(\theta, x) \, dx \tag{1.25}$$

where $n(x)$ is the size distribution, and $\sigma(\theta, x)$ is the volume scattering function for a particles of radius x, neglecting the scattering from index of refraction fluctutations and Rayleigh scattering from molecules.

The dependence of the volume scattering function $\beta(\theta)$ on size distribution is a good measure of importance of the size distribution. Figure 1.1,

[1]The convention for optical oceanographers now is to use c instead of α for the attenuation coefficient [1]. I continue to use α, to distinguish it from the constant c, a parameter in the Junge distribution, to be discussed in section 1.3.

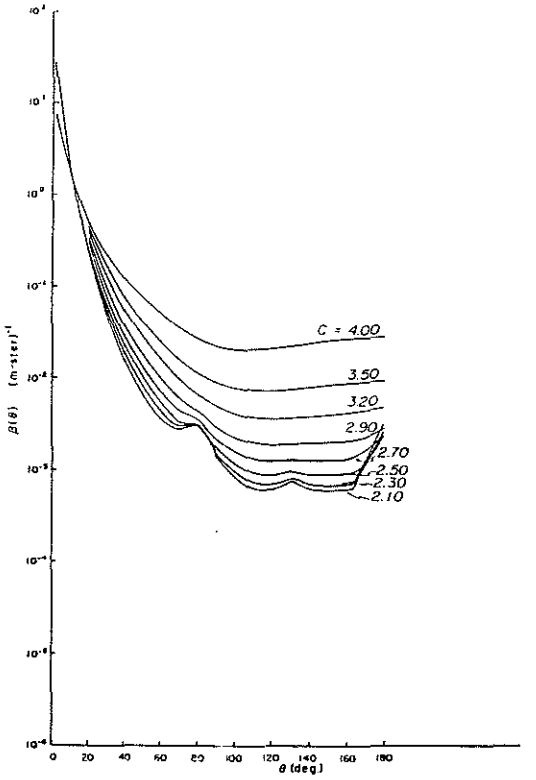

Figure 1.1: $\beta(\theta)$ for different $n(x)$, computed by Morel, for various constants c. From Gordon *et al.* [1].

from [1] illustrates the dependence of $\beta(\theta)$ on size distribution. There, $\beta(\theta)$ is plotted for a series of so called Junge distributions (see section 1.3) where the parameter c is varied. The effect of size distribution on $\beta(\theta)$ is clearly evident. The size distributions with lower c, which correspond to distributions with relatively greater concentrations of large particles scatter mostly into small angles, as is expected. On the other hand, distributions skewed toward the small sizes scatter mostly into large angles.

The effect of particle size distribution in affecting underwater visibility is evident in the relation between the modulation transfer function and the volume scattering function. The modulation transfer function is the spatial frequency response function of image propagation, and is familiar in the study of optical systems. The relation between the modulation transfer

function and the volume scattering function was first pointed out by Wells [22], and investigated experimentally by Del Grosso [23]:

$$\gamma(\psi, R) = e^{\left[\alpha - \int_0^{\theta_n} \beta(\theta) J_0(2\pi\theta\psi)\theta \, d\theta\right]} \tag{1.26}$$

where $\gamma(\psi, R)$ is the transfer function, ψ is spatial frequency, and R is the range of propagation. J_0 is a Bessel function, and θ is the scattering angle. The upper limit of integration, θ_n requires judicious choice, but can be quite small since the transfer function is dominated by scattering at small θ.

1.3 Survey of Size Distributions in the Ocean

Despite the importance of particle size distributions in physical and chemical phenomena in the ocean, measurements have been few and incomplete. In the literature, most measurements of size distributions come from the deep ocean; surprisingly, few come from coastal waters. Size distributions vary widely; some deep ocean water is "cleaner" than tap water, while some coastal water is turbid. Most of the data is obtained using Coulter counters. Unfortunately, the lower limit diameter for Coulter count measurements is about 1 μm and, as McCave notes, however, most particles in the ocean have diameter less than 2 μm [2].

Due to the limitations of Coulter counters, oceanic size distributions are best documented for particles with diameter larger than one micron. For such distributions the *Junge* distribution, a negative power law has been popular [24]:

$$N(D) = kD^{-c} \tag{1.27}$$

where $N(D)$ is the *cumulative* number density of particles, of diameter D. For non-spherical particles, D is chosen to be the same as particles of equivalent volume or cross section.[2] This corresponds to a number density

$$n(D) = -\frac{dN(D)}{dD} = ckD^{-(c+1)} \tag{1.28}$$

Carder observed that the Junge distribution is

> a good first approximation (and often better than that) descriptive of many natural families of particles such as fine sediments, airborne dust, cosmic dust, suspended particles in sea water, lunar surface material and even very large particles producing lunar craters by impact [25].

A few investigators have fit the Junge distribution to measurements from the ocean, obtaining measurements with Coulter counters, mostly. Their results are summarized in table 1.1. Typically, as the table shows, k varies substantially, from about 10^3 to 10^5 or more per ml, while c is approximately 3 universally. (In table 1.1, Baker's data comes from scanning photomicrographs; the Junge distribution is fit only for the larger particles; otherwise, it would appear that the log-normal distribution, to be discussed later, would fit better [26].)

While the Junge distribution fits well for D greater than 1 μm, it does not fit for smaller particles. In fact, it cannot fit for $D \to 0$, because equation (1.27) becomes singular.

[2] In this work I will usually refer to number density $n(x)$, where $x = kr$ is wave-number non-dimensionalized radius. Our notation in this section, however, is consistant with the literature on oceanic particles.

Investigator	Location	Size range, μm	k /ml	c	Comments
Sheldon (1967) [27]	Sannich Inlet	1 - 100		≈ 3	British Columbia
Bader (1970) [28]	Bahamas, Abaco Bight	2 - 24	$\approx$20000	1.77 - 2.4	Wind Agitated Coastal water
			29,845	2.64	D < 4.6 μm
			466,000	4.36	D > 4.6 μm
					Same, "quiet"
Sheldon *et al.* (1972) [29]	Various deep ocean waters	1 - 100		≈ 3	
Brown, Gordon (1973) [30]	Sargasso Sea	0.08 - 10	$11x10^3$	3	
McCave (1975) [2]	(from Sheldon)	1 - 100	10^3-10^4	2.4 - 3.6	"clear" water
Lal, Lerman (1975) [31]	South Indian Ocean	1 - 100		3.75 - 4.5	Foraminifera and diatom shells only
Bishop *et al.* (1978) [32]	Southeast Atlantic	>50	$10^4 - 10^5$	4.6 - 7.7	Foraminifera only
			$10^3 - 10^7$	4.1 - 6.3	Foraminifera Fragments
			$10^4 - 10^5$	2.8 - 4.6	Fecal Pellets
			$10^4 - 10^9$	3.8 - 4.9	Fecal Matter
Baker (1979) [26]	Eastern Equatorial Pacific	0.65 - 4	300 - 1000	3.29$\pm$.3	mid water in 5000m
Pak (1980) [33]			2.9 - 3.5		
Spinrad *et al.* (1983) [34]	Nova Scotian Rise			3.1 ± 0.3	HEBBLE experiment
Tsuda, Nakata (1982) [35]	Akita Bay, Japan	2.3 - 36.2	44,000 - 300,000	$\approx$2.5	Coastal water
McCave (1983) [2]	Nova Scotian Rise			2.6– 3.2	HEBBLE experiment

Table 1.1: Summary of Data to which Junge Distribution has been fitted.

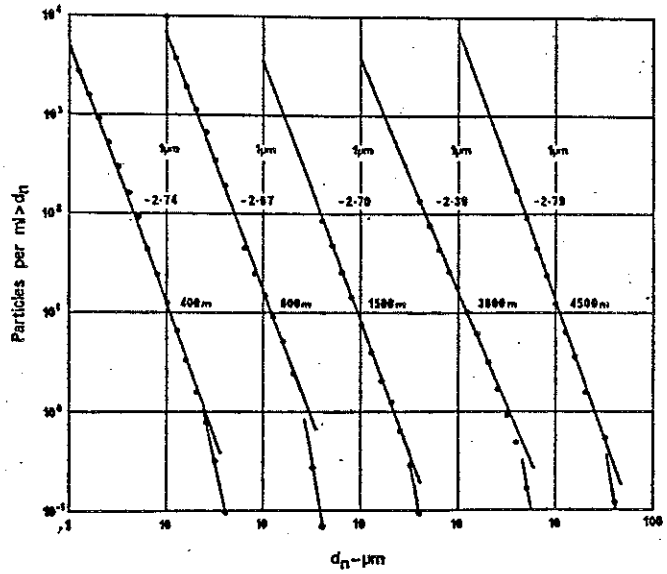

Figure 1.2: Cumulative number densities of particles at different depths, measured by Coulter counter, from data by Sheldon. Slopes of distributions are shown. Figure from McCave [2].

The Junge distribution plots as a straight line on a log-log plot; the constant c in equation (1.27) is the negative of the slope. That the Junge distribution does not fit for small particles is generally not evident in Coulter count measurements. See for example, figure 1.2.

The failure of the Junge distribution is especially evident in the GEOSECS data of Lambert *et al.* [3] from the western Atlantic, in which size distributions of alumino silicate particles as small as 0.2 μm were obtained by counting under scanning electron microscope [3]. Figure 1.3 is a log-log plot of a size distribution, typical of their data.

In figure 1.3, $N(D)$ is a straight line for D greater than 1 μm; above

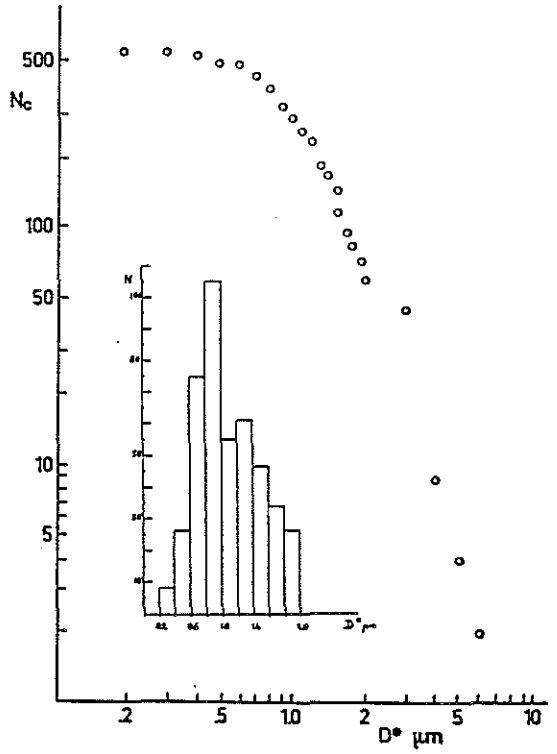

Figure 1.3: Cumulative size distribution of Alumino-silicate particles. From Lambert *et al.* [3]. The inset plot is number density.

1 μm the Junge distribution works well.[3] For D less than 1 μm, however, where Coulter count data were not available, the curve flattens, showing that the Junge distribution breaks down.

Lambert *et al.* are among the few investigators who fit data to a distribution other than a negative power law. A log-normal distribution fits their data for small particles as well as for large particles:

$$n(D) = \frac{N_t}{\sigma\sqrt{2\pi}D^*}\exp(-u^2/2) \tag{1.29}$$

where

$$u = \frac{\ln D - \ln D^*}{\sigma} \tag{1.30}$$

Here, σ is the standard deviation of the distribution, N_t is the total number

[3]The constant c is the negative of the slope of this curve, on a log-log plot.

of particles in the distribution, and D^* is the average size in the distribution. For large distributions, the power law and log normal distributions are indistinguishable.

For GEOSECS data, Lambert *et al.* D^* is typically 1.5 to 2.0 μm, while σ is typically from 0.5 to 0.7. These results lack generality, however, since they apply only to the alumino silicate particles present. For coastal waters, for example, D^* may be larger.

The Junge distribution has been used in some light scattering computations, but these usually depend on patching the distribution, such as estimating a lower cut-off for the distribution, for example [36,30,37,38,35,39]. Few investigators of oceanic optics have used the log-normal distribution. Pak *et al.* have utilized a log-normal distribution, for computing Mie scattering from suspended particles, of a slightly different form from equation (1.30) [33]:

$$\nu(D) = \frac{N_t}{D\sigma\sqrt{2\pi}} \exp -u^2/2 \tag{1.31}$$

where $\nu(D)$ is the *volume* distribution, related to number density $n(d)$ as follows:

$$n(D) = \frac{6\nu(D)}{\pi D^3} \tag{1.32}$$

They report that for data taken by Revelle appropriate values of μ and σ^2 are 1.5 μm and 15 μm^2 respectively.

Ensor and Pilat have also utilized a log-normal distribution in their work, of the form of equation (1.30) for scattering computations, but regrettably present no experimental measurements of parameters used [40].

The lack of data regarding size distributions in the ocean, especially at

small sizes, makes it difficult to predict light scattering from oceanographic particles. This poses a problem for the designer of an instrument to measure scattering with the intention of inverting for size distributions. It is difficult to design an appropriate detector in advance without knowing something about the scattering to be measured. In part, this drives my experimental program. One goal of the program is to measure the scattering from natural distributions, so as to determine the necessary sampling rate and interval.

1.4 Survey of Particle Sizing Methods

In this section I review briefly methods of determining particle size distributions other than diffraction sizing. This review is for comparison; for more comprehensive reviews I refer the reader to the literature [41,42]. The methods are distinguished by the size range of particles they cover, and whether or not they are intrusive.

Scanning under optical and electron microscopes is the most direct sizing method. This has the advantages of sizing particles of nearly any size, and distinguishing between types of particles. Lambert *et al.* [3], for example, sized alumino-silicate particles as small as 0.2 μm under a scanning electron microscope. The method is, however, time consuming, and destroys the sample under study.

Unlike scanning under microscopes, modern imaging methods need not be intrusive, as, for example, holographic imaging [43]. Holographic imaging has the advantage that particle velocities can be deduced from a time series of images. Carder *et al.* [44], for example used holography *in situ* for

measuring sizes and settling rates of particles in the western North Atlantic ocean; the smallest size they could measure was 15 μm. The non-intrusive methods lack ability to size particle smaller than about 5 μm, thus missing most particles, and require tedious or computationally intensive processing.

There are several optical sizing methods, besides imaging, including diffraction sizing. One such method, which has been pursued for oceanographic work, is the photo-extinction method. In this method, the optical extinction of a closed sample is measured. As particles settle out, optical extinction decreases. By assuming an optical extinction function as a function of particle radius and Stoke's law for the settling rate, a particle size distribution can be deduced. The maximum particle size for this method is determined by the validity of Stoke's law. Stoke's law is valid when the Reynolds number is less than 0.5; particles less than 100 μm qualify easily. Zanefeld *et. al.* [45] reported on an *in situ* instrument for the ocean which measures as small as 2 μm, depending on the concentration of the smallest particles. Since the method relies on holding a sample volume captive, however, it is necessarily intrusive. The time for analyzing a sample is long; approximately 24 hours. Furthermore, since the method relies on Stoke's law, it can not be used to measure settling velocities.

Another set of optical methods are based on laser doppler velocimetry [46,47,48]. These methods count single particles, analyzing the doppler return for size as well as velocity. In one method, absolute scattered intensity is correlated with the scattering computed from any of several scattering theories. In another, the frequency content of the return is analyzed for size. These methods are non-intrusive. Ariessohn *et al.* were able to measure

particles of coal as small as 0.2 μm [47]. However, probe volumes are particularly small, so that measurement times must be long to get statistically significant results. Furthermore, the instrumentation is complicated.

Of all particle sizing methods, Coulter counting is best known. Most measurement of size distributions in the ocean, as shown in section 1.3, have been made via Coulter counters. The principle of the method is that as a particle passes through an aperture, it changes the electrical properties of the aperture by an amount proportional to the volume of the particle. In practice the smallest size measurable is about 1 μm diameter. The method is necessarily intrusive, since the fluid must pass through an aperture.

1.5 Diffraction Particle Sizing: The Method

Diffraction particle sizing is an out-growth of optical computing, or "Fourier optics" [49,50]. It was first described by Chin, Sliepcevich and Tribus [8,9]; some later work is listed in the references [51,52,53,54,55,56,57,58,59,60,61].

The technique is based on numerically inverting measurements of the angular distribution of near forward scattering of a collimated, coherent beam by a suspension of particles. A simplified schematic of a diffraction particle sizing instrument is shown in figure 1.4. A laser beam, filtered and expanded through a spatial filter, passes through a sample volume. The scattered field (and the unscattered beam) are focused by a lens onto a multi-element detector.

The effect of the lens, according to ray theory, is to sort the scattered rays by angle, so that all rays with an angle θ impinge on the detector at

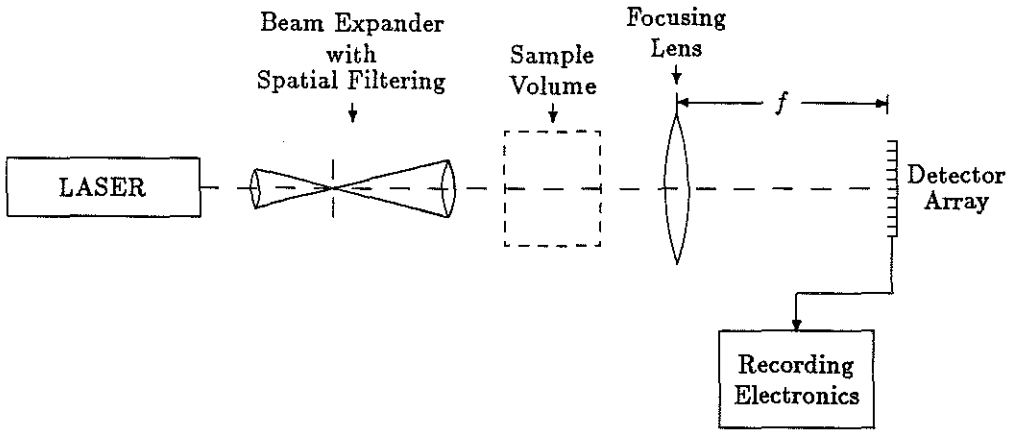

Figure 1.4: Simplified schematic of a diffraction particle sizing instrument.

a distance l off center, given by

$$\theta = \frac{l}{f} \tag{1.33}$$

where f is the focal length of the lens. This is shown in figure 1.5.

A small angle approximation $\sin\theta \approx \theta$ is implicit. The advantage of this configuration is that, barring effects of vignetting by stops and lenses, the intensity on the detector is independent of the positions of the particles. This simplifies the inversion, though the extreme magnitude of the unscattered beam near $\theta = 0$ becomes an experimental problem.

Alternately, the lens can be viewed as a Fourier transformer. Then, according to the diffraction approximation of section 2.2, particles are approximately apertures with the same geometric shadows. The image on the detector is the sum of the Fourier transforms of these apertures. The positions of particles are contained purely in the *phases* of their Fourier

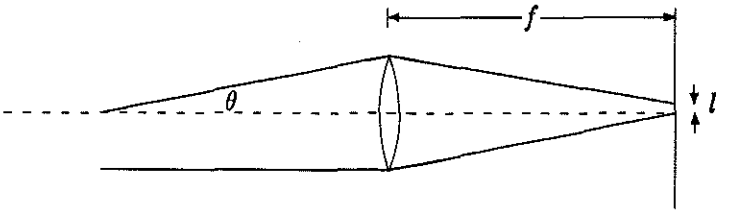

Figure 1.5: Effect of lens, according to a ray theory argument. $\theta = 0$ for the ray at the bottom.

transforms. The phases are discarded by the detector which observes intensity (magnitude) only, thus removing the complication of particle position. I will elaborate on the use of a lens as a Fourier transformer in section 2.8.

The detector measures an angular intensity distribution, which is a sum of the scattered intensity, $I(\theta)$, which is to be inverted to yield the size distribution, and the unscattered beam, $I_d(\theta)$. In principle, $I_d(\theta)$ is concentrated entirely at $\theta = 0$, so $I(\theta)$ and $I_d(\theta)$ are separable by position.

The output of the detector, $I(\theta)$, feeds to a computer for inversion. The process of inverting $I(\theta)$ to yield the size distribution $n(x)$ will be the topic of chapter 4.

Diffraction particle sizing is potentially non-intrusive, can be used to obtain time series of size distribution, and has a lower limit particle size of at least a wavelength. Of the methods discussed, this alone promises measurements over a wide range of sizes, instantaneously and repeatedly, and non-intrusively.

Chapter 2

Scattering Theories

As a precursor to the theoretical issues of inverting the scattered intensity $I(\theta)$ to yield the size distribution $n(x)$, which I call the *inverse* problem, I discuss the *forward* problem, finding $I(\theta)$ from $n(x)$. This is the underlying physics of diffraction particle sizing — the scattering from small particles.

In this thesis I often refer to an "intensity", denoted by I. This is often called *irradiance*, and refers to the total radiant power falling on a detector of area dA, divided by the area:

$$I = \frac{d\phi}{dA}, \text{ watts}/m^2$$

where ϕ is the radiant flux:

$$\phi = \frac{d(\text{energy transported})}{d(\text{time})}$$

It is implicit that the detector is perpendicular to the incident beam. The argument of the intensity always refers to a position in space; thus when I refer to an intensity at an angle θ, e.g. $I(\theta)$, a range R is implicit. Furthermore, for a scattered intensity $I(\theta)$ from a distribution of scatters,

the intensity is implicitly per unit volume. Then, scattered intensity $I(\theta)$ corresponds to a volume scattering function

$$\beta(\theta) = R^2 I(\theta) \tag{2.1}$$

where R is the distance from the scatterer.

I denote the scattered intensity at angle θ from a particle of wavenumber non-dimensionalized size $x = kr$, by $I(\theta, x, \eta)$, where η is the complex index of refraction. In the next sections, I discuss pertinent models for $I(\theta, x, \eta)$; then, I model $I(\theta)$ for a distribution of particles of refractive index η, making appropriate simplifying assumptions.

I assume that the scattering from a suspension of particles is due entirely to the particles themselves, and neglect scattering from other sources, such as the Rayleigh scattering from water molecules or variations of index of refraction. In the ocean, the scattering from particles dominates.

In this work, three models for $I(\theta, x, \eta)$ are appropriate — the exact solution of the wave equation for spheres, known as Mie scattering theory, and two approximations, the diffraction approximation and Rayleigh scattering.

For each of the three models the incident field consists of plane waves. This assumption is good even for the Gaussian field of a laser, because the particles are small relative to the beam radius.

2.1 Mie Scattering

For arbitrary particles, the exact scattering $I(\theta, x, \eta)$ is the solution of Maxwell's equations with appropriate boundary conditions. When applied to spheres, the solution is known as Mie scattering. I leave the lengthy derivation to the literature [62,63,64,65]. Following the notation of van de Hulst [64], the far-field scattering from a sphere of radius x is written as:

$$S_1(\theta) = \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} \{a_n \pi_n(\cos\theta) + b_n \tau_n(\cos\theta)\} \tag{2.2}$$

$$S_2(\theta) = \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} \{b_n \pi_n(\cos\theta) + a_n \tau_n(\cos\theta)\} \tag{2.3}$$

where the functions $\pi_n(\cos\theta)$ and $\tau_n(\cos\theta)$ are written in terms of the Legendre polynomials:

$$\pi_n(\cos\theta) = \frac{1}{\sin\theta} P_n^1(\cos\theta) \tag{2.4}$$

$$\tau_n(\cos\theta) = \frac{d}{d\theta} P_n^1(\cos\theta) \tag{2.5}$$

The coefficients a_n and b_n are

$$a_n = \frac{\psi_n'(y)\psi_n(x) - \eta\psi_n(y)\psi_n'(x)}{\psi_n'(y)\varsigma_n(x) - \eta\psi_n(y)\varsigma'(x)} \tag{2.6}$$

and

$$b_n = \frac{\eta\psi_n'(y)\psi_n(x) - \psi_n(y)\psi_n'(x)}{\eta\psi_n'(y)\varsigma_n(x) - \psi_n(y)\varsigma'(x)} \tag{2.7}$$

with

$$\psi_n(y) = y j_n(y) = \left(\frac{\pi y}{2}\right)^{1/2} J_{n+1/2}(y) \tag{2.8}$$

and

$$\varsigma_n(y) = y h_n^{(2)}(y) = \left(\frac{\pi x}{2}\right)^{1/2} H_{n+1/2}^{(2)}(y) \tag{2.9}$$

Here, $y = \eta x$, and $j_n(y)$ and $h_n^{(2)}(y)$ are the spherical Bessel functions written in terms of the fractional order Bessel functions $J_{n+1/2}(y)$, and $Y_{n+1/2}(y)$, respectively. Furthermore, η is a complex quantity, $\eta = \eta' + i\eta''$, where the complex part corresponds to an absorption coefficient.

The scattered intensities are, in units of watts/m^2,

$$i_1(\theta) = \frac{|S_1(\theta)|^2}{k^2 R^2} \tag{2.10}$$

for incident "perpendicular" polarization, for which the electric field is perpendicular to the plane of scattering, and

$$i_2(\theta) = \frac{|S_2(\theta)|^2}{k^2 R^2} \tag{2.11}$$

for incident "parallel" polarization, for which the electric field is parallel to the plane of scattering. R is the range from the scatterer.

Figures 2.1, 2.2 and 2.3 show examples of $I(\theta, x, \eta)$ computed from Mie theory, for three sizes of particles. Note that $i_1(\theta)$ and $i_2(\theta)$ are indistinguishable at small θ. Thenceforth, when considering scattering at small angles, I will not draw any distinctions between the two polarizations.

Unfortunately, the equations of Mie scattering are difficult to manipulate and are time consuming to compute [66,67,68,69,70]. Thus, I use the diffraction and Rayleigh approximations exclusively in the theoretical work.

2.2 The Diffraction Approximation

For the theoretical work, diffraction theory yields a simple model of scattering which is easily manipulated and is physically robust. Since I use

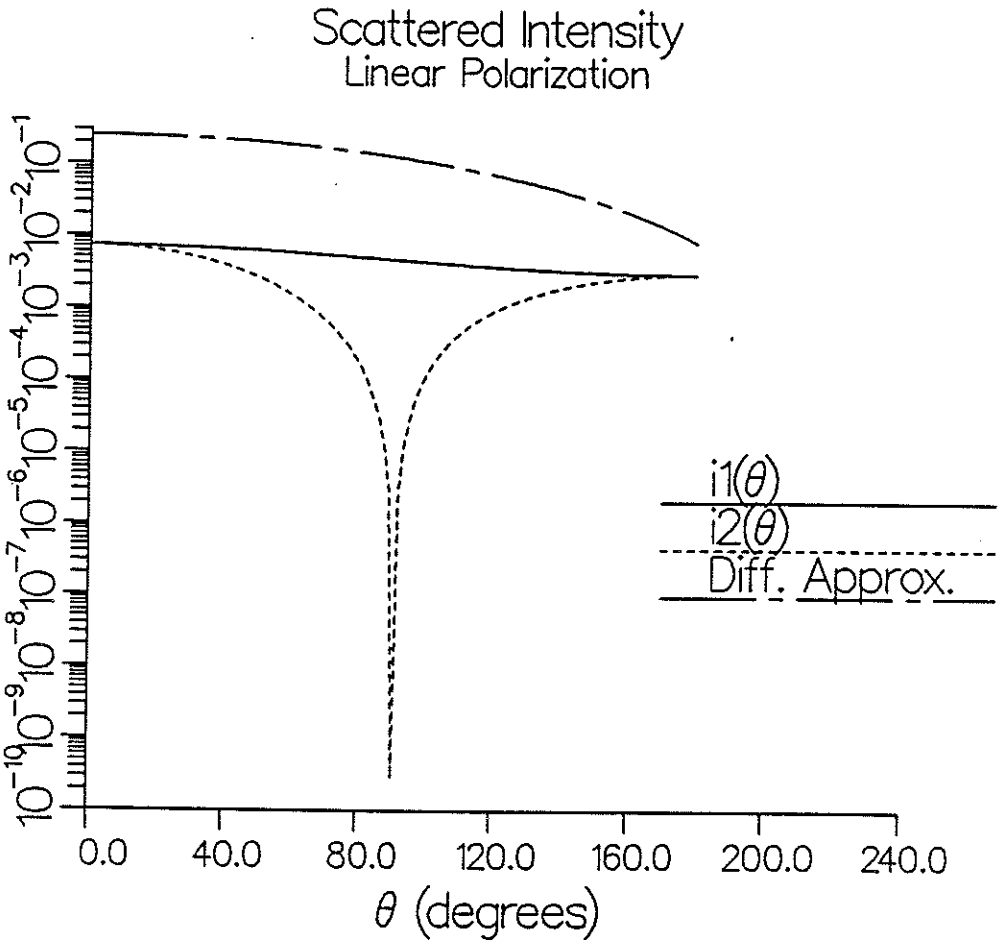

Figure 2.1: Mie scattering for a sphere of radius $x = 1$ and $\eta = 1.125 - i0.001$. The top curve represents the diffraction approximation, but the Rayleigh approximation is most appropriate for this size particle.

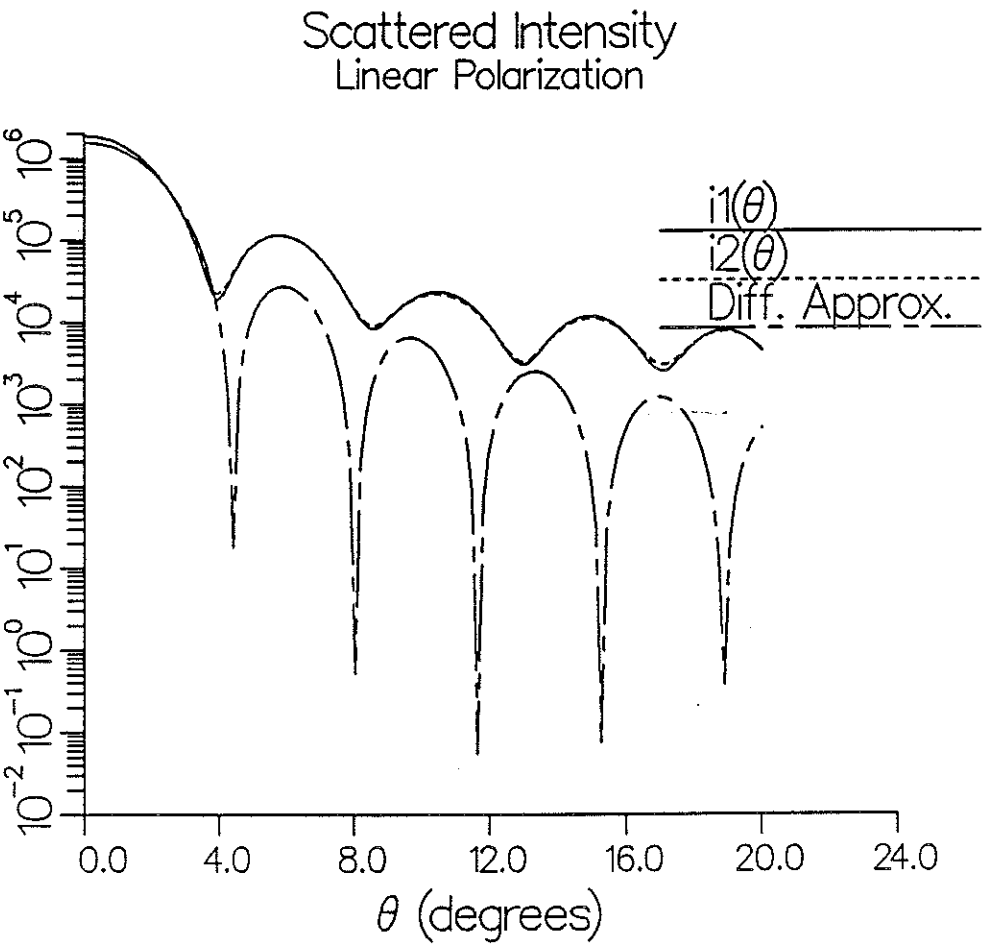

Figure 2.2: Mie scattering for a sphere of $x = 50$ and $\eta = 1.125 - i0.001$. For comparison, I include the diffraction approximation.

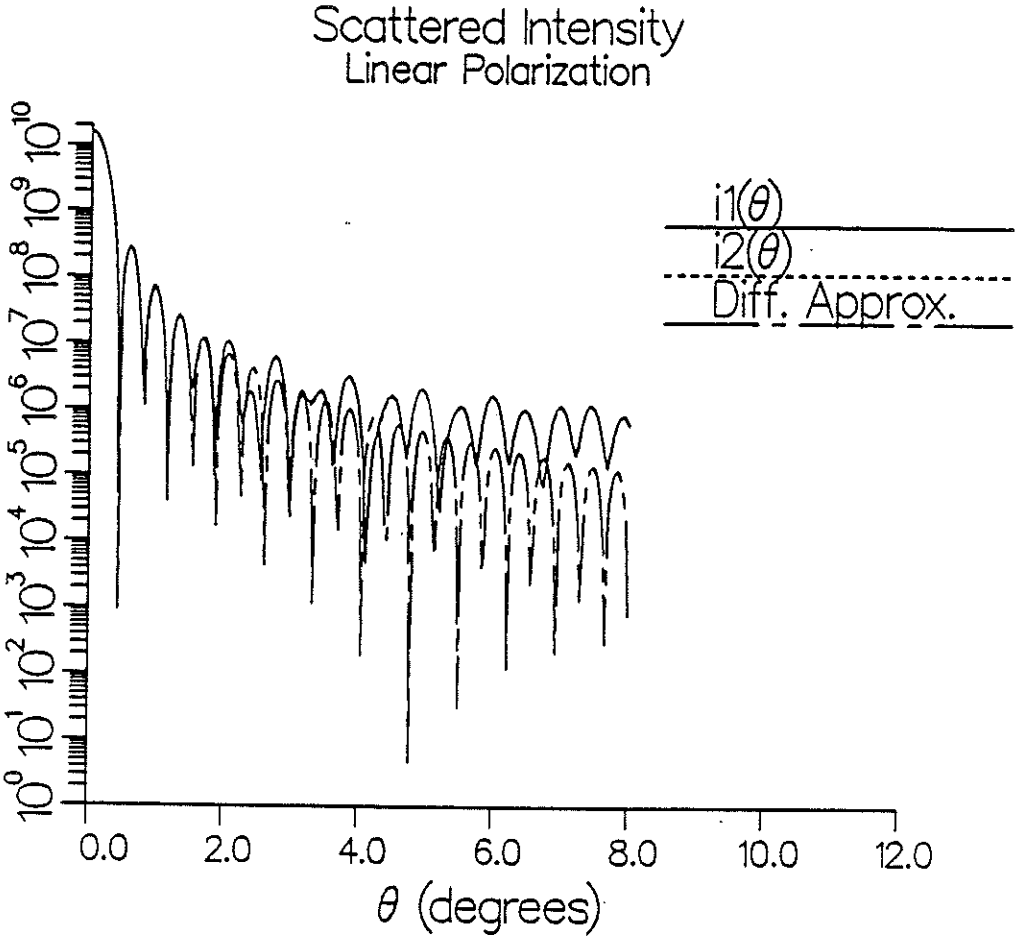

Figure 2.3: Mie scattering for a sphere of $x = 500$ and $\eta = 1.125 - i0.001$. For comparison, I include the diffraction approximation.

the diffraction approximation repeatedly, I derive it in the paragraphs that follow. The derivation follows Haus [71]; Gaskill offers a similar derivation [49]. The derivation starts with the wave equation for the electric field $\mathbf{E}(\mathbf{r},t)$ in a homogeneous medium:

$$\nabla^2\mathbf{E}(\mathbf{r},t) - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\mathbf{E}(\mathbf{r},t) = 0 \tag{2.12}$$

Here $\mathbf{r} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}}$, where $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$ are the unit vectors running along the x, y, z axes, respectively. Now, assuming a polarization and time dependence,

$$\mathbf{E}(\mathbf{r},t) = \hat{\mathbf{x}}u(\mathbf{r})e^{-i\omega t}, \tag{2.13}$$

where $u(\mathbf{r})$ is the scalar field amplitude, one obtainds the time independent Helmholtz equation:

$$\nabla^2 u(\mathbf{r}) + k^2 u(\mathbf{r}) = 0 \tag{2.14}$$

where $k = \omega/c$, and $c = \eta c_0$, η is the (real) index of refraction, and c_0 is the speed of light in a vacuum.

The Helmholtz equation admits solutions of the form $e^{i\mathbf{k}\cdot\mathbf{r}}$, which corresponds to plane waves, from which one can build other, arbitrary solutions by integrating over wave-number:

$$u(\mathbf{r}) = \int_{-\infty}^{\infty} U(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{r}}d\mathbf{k} \tag{2.15}$$

Here $\mathbf{k} = k_x\hat{\mathbf{x}} + k_y\hat{\mathbf{y}} + k_z\hat{\mathbf{z}}$.

I assume a field which propagates nearly on-axis, so that

$$k_z = \sqrt{k^2 - k_x^2 - k_y^2} \approx k - \frac{k_x^2 + k_y^2}{2k} \tag{2.16}$$

and limit the discussion so that this is valid, i.e. so that $U(k_x, k_y)$ is non-zero only when

$$\frac{\sqrt{k_x^2 + k_y^2}}{k} \ll 1 \tag{2.17}$$

Equation (2.15) becomes

$$u(\mathbf{r}) = \int_{-\infty}^{\infty} dk_x \int_{-\infty}^{\infty} dk_y \, U(k_x, k_y) e^{-i(k_x x + k_y y)} e^{i\frac{(k_x^2 + k_y^2)}{2k} z} \tag{2.18}$$

where $U(k_x, k_y)$ is a complex amplitude, a function of the remaining wave-numbers k_x and k_y.

One can see how a field can be propagated from one range to another by setting $z = 0$ in equation (2.18). The result is that equation (2.18) is a Fourier transform, which has the inverse transform

$$U_0(k_x, k_y) = \left(\frac{1}{2\pi}\right)^2 \int_{-\infty}^{\infty} dx_0 \int_{-\infty}^{\infty} dy_0 \, u_0(x_0, y_0) e^{i(k_x x_0 + k_y y_0)} \tag{2.19}$$

Here I use $u_0(x_0, y_0)$ to represent $u(x, y, 0)$, which has the Fourier transform $U_0(k_x, k_y)$.

Equation (2.19) is now substituted back into equation (2.18). Some algebra yields the Fresnel diffraction integral, which propagates a field with complex amplitude $u_0(x_0, y_0)$ at $z = 0$, to $u(x, y)$ at a range z:

$$u(x, y, z) = \frac{ik}{2\pi z} \int_{-\infty}^{\infty} dx_0 \int_{-\infty}^{\infty} dy_0 \, u_0(x_0, y_0) e^{-i\frac{k}{2z}[(x - x_0)^2 + (y - y_0)^2]} \tag{2.20}$$

The next result is the Fraunhoffer diffraction integral, the far-field limit of (2.20). In this limit, the following substitution is made in equation (2.20):

$$\frac{k}{z}\left[(x - x_0)^2 + (y - y_0)^2\right] \approx \frac{k}{z}\left[(x^2 + z^2) - 2xx_0 - 2yy_0\right] \tag{2.21}$$

where the assumption is made that the missing term, $k(x_0^2+y_0^2)/z$ is small, i.e. that

$$z \gg k(x_0^2+y_0^2) \tag{2.22}$$

The Fraunhoffer approximation is then

$$u(x,y,z) = \frac{ik}{2\pi z} e^{-i\frac{k(x^2+y^2)}{2z}} \int_{-\infty}^{\infty} dx_0 \int_{-\infty}^{\infty} dy_0\, u_0(x_0,u_0) e^{\frac{ik}{z}(xx_0+yy_0)} \tag{2.23}$$

The field at large range is, to within a complex phase factor, a Fourier transform of the field at $z=0$.

I make several observations regarding equation (2.23), which are relevant to later sections:

- The intensity $I(\mathbf{r})$ is given by $I(\mathbf{r}) = \langle u(\mathbf{r})u^*(\mathbf{r})\rangle$, where brackets denote ensemble averaging.

- The Fourier spectrum of $u(x,y,z)$ is $u_0(x_0,u_0)$, within a constant phase factor. Thus, $u(x,y,z)$ is *band-limited* if $u_0(x_0,y_0) = 0$ for $x_0^2+y_0^2 > D$, where D is an upper bound.

- If $u(x,y,z)$ is band-limited, then so is the intensity $I(\mathbf{r})$.

In the diffraction approximation, the scattering from a particle is modelled as the Fraunhoffer diffraction from an aperture with the same geometric shadow. The rationale is that if a particle is sufficiently refracting or attenuating, some of the incident wave is "lost" by refraction into large angles or by attenuation. The remaining wave is the same as that of an aperture of equivalent geometric shadow. If the aperture is large, the diffracted

component is concentrated in small angles and dominates the refracted component. Thus the region of validity of the diffraction approximation is

$$x \gg 1 \tag{2.24}$$

$$|x(\eta - 1)| \gg 1 \tag{2.25}$$

$$\theta \ll 1 \tag{2.26}$$

where the inequality in equation (2.25) means that the phase change of the diffracted rays must be large. When $\eta \approx 1$, that is, the particle is weakly refracting, the refracted and diffracted components of scattering are comparable in strength and interfere. In this case the theory of *anomalous diffraction* is appropriate.

The diffraction approximation is a polarization independent theory. Also, except for its region of validity, it is independent of refractive index.

For spheres the diffraction approximation is the same as scattering from a disk of the same radius, so that substituting

$$u(x_0, y_0) = \begin{cases} 1 & \text{if } \sqrt{x_0^2 + y_0^2} < x \\ 0 & \text{otherwise} \end{cases} \tag{2.27}$$

into equation (2.23) gives

$$I(\theta, x) = \frac{x^4}{k^2 z^2} \frac{J_1^2(x\theta)}{(x\theta)^2} \tag{2.28}$$

where I assume unit incident intensity in units of watts/m^2, and J_1 is the first order Bessel function of the first kind. Here I introduce the scattering angle

$$\theta = \tan^{-1}\left(\frac{\sqrt{x^2 + y^2}}{z}\right) \approx \frac{\sqrt{x^2 + y^2}}{z}, \tag{2.29}$$

and make the approximation $\sin\theta \approx \theta$.

I use the diffraction approximation for spheres frequently in this work. I use it for particles other than spheres as well, assuming that a distribution of odd shaped particles are statistically spherical in the sense that for the diffraction approximation, an equivalent distribution of spherical particles can be found that have the same scattering.

The diffraction approximation is better than the region of validity set by equations (2.24), (2.25) and (2.26) suggests. Bayvel and Jones [63] report that for transparent spheres the difference between Mie theory and the diffraction approximation is less than 1% for $\eta > 1.3$ and $x > 20$. For absorbing particles with $\eta = 1.2 - i0.1$ the difference is less than 20% for $x > 3$. Other investigators who have investigated the region of validity of the diffraction approximation are Jones [72] and Boron and Waldie [73]. The region of validity of the diffraction approximation is also evident in figures 2.1, 2.2, and 2.3.

I use the diffraction approximation to primarily assess the frequency content and asymptotic behavior of the scattering $I(\theta)$ from an assembly of particles. For this purpose, it is not so important that the diffraction approximation is accurate, as that it predicts the frequency content and asymptotic behavior accurately. For inverting experimental data with the approximation, however, the accuracy of the approximation *is* important. This is particularly true for inversion algorithms based on the Chin-Sliepcevich integral transform [74,75,76,11].

2.3 Rayleigh Scattering

Rayleigh scattering applies to the smallest particles, for $x \ll 1$ and $x(\eta - 1) \ll 1$. Such particles "see" an almost homogeneous, though time varying field, and oscillate like a simple dipole with polarization proportional to the incident field. The field from such particles can be written in matrix form [63]:

$$\begin{pmatrix} E_{\parallel sca} \\ E_{\perp sca} \end{pmatrix} = \frac{e^{ikD}}{D} k^2 \nu \begin{pmatrix} \cos\theta & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} E_{\parallel inc} \\ E_{\perp inc} \end{pmatrix} \tag{2.30}$$

where the subscripts $\parallel$ and $\perp$ correspond the parallel and perpendicular polarizations, respectively, and *sca* and *inc* refer to the scattered and incident fields. The parameter ν is the particle polarizability, and depends on shape and index of refraction. From this the scattered intensity can be written

$$\begin{pmatrix} I_{\parallel sca}(\theta) \\ I_{\perp sca}(\theta) \end{pmatrix} = \frac{k^4 |\nu|^2}{D^2} \begin{pmatrix} I_{\parallel inc} \cos^2\theta \\ I_{\perp inc} \end{pmatrix} \tag{2.31}$$

For spheres, the parameter ν is

$$\nu = \frac{x^3}{k^3} \frac{\eta^2 - 1}{\eta^2 + 2} \tag{2.32}$$

where η is the index of refraction. ν can be calculated for other particles, but for non-symmetrical particles, depends on orientation.

2.4 Distributions of Particles

In this work, several simplifying assumptions are appropriate regarding scattering from dispersions of particles:

1. The particles are sufficiently separated that they scatter independently.

2. The optical depth of the scattering volume is small, so that the incident field is not attenuated.

3. The particles are positioned and oriented randomly, so that no Bragg scattering occurs.

4. There are enough particles present that the size distribution $n(x)$ can be represented as a continuum.

Under these assumptions single scattering is said to occur, and the first Born approximation applies. The scattering from a suspension is given by the integral

$$I(\theta) = \int_0^\infty n(x) I(\theta, x)\, dx \tag{2.33}$$

Were the fourth assumption false, the distribution could be modeled as a discrete sum over individual particles. In fact, in section 2.7 I show that equation (2.33) actually represents an ensemble average of a stochastic quantity.

To justify equation (2.33) one must determine the region of validity of simplifying assumptions 1 and 2.

Assumption 1 From detailed computations of electromagnetic fields, the interactions between particles appear to be negligible if the particles are separated by at least 3 diameters [63,77,65]. However, Napper and Ottewill found in measurements using a polystyrene latex dispersion that multiple

scattering effects are present unless the distance between particles is greater than 100 diameters. In either case, the separations correspond to large concentrations.

Assumption 2 The optical depth for which single scattering is valid may be estimated with the following argument. A photon travels through a suspension, encountering scatterers with average extinction cross-section σ_{ext}, with a mean free path l between collisions. So that l is the mean free path, there must be only one particle, on average, in the volume $\sigma_{ext}l$. Thus, the number of particles per unit volume, N, must be

$$N = \frac{1}{\sigma_{ext}l} \tag{2.34}$$

The mean free path is

$$l = \frac{1}{N\sigma_{ext}} \tag{2.35}$$

The quantity $N\sigma_{ext}$ can be identified as the scattering coefficient b in section 1.2.2. Multiple scattering occurs when the optical depth, L, is large; compared to the mean free path, when $L \gg l$. Single scattering occurs if $L \ll 1/b$. Bayvel suggests the criterion

$$L \ll \frac{0.1}{b} \tag{2.36}$$

In this work, I assume that these assumptions are valid, and justify them experimentally. If single scattering is valid, then $I(\theta)$ has a linear dependence on particle concentration. If additional particles are added to a suspension, $I(\theta)$ increases by a purely multiplicative factor, the relative increase in concentration. Alternatively, if the length of a scattering volume

could be varied, increasing the length would result in a purely exponential decrease in $I(\theta)$ for all θ.

An example of how I verify that single scattering is a good approximation in experiments is shown in figure 2.4. To produce the data for this figure, I measured the light scattering of seven different concentrations of polystyrene spheres, with a narrow Gaussian size distribution with mean diameter 40.3μm. Each curve represents a scattered intensity, divided by the corresponding concentration of spheres. Details regarding the spheres are summarized in table 6.1. The concentrations were produced by adding drops of a standard solution to the measurement volume. The measurement volume is one inch thick, with a volume of 70ml. Further experimental details are contained in a later chapter.

The curves in figure 2.4 overlay quite closely, indicating that the dependence of scattering on concentrations at these levels is purely multiplicative. This suggests that single scattering is indeed an excellent approximation. The significant differences between curves in figure 2.4 can be explained entirely by experimental noise, especially by digitization noise for low intensities and low concentrations of particles. Identifying symbols on the curves are omitted for clarity; however, the curves for the higher concentrations start at larger angles due to saturation of photodiodes.

For the thin sample volume I use, single scattering is valid even for very high concentrations of particles. For comparison, table 2.1 lists the 7 concentrations, corresponding to the curves in figure 2.4, with estimates for τ and attenuation coefficient α. (The procedure for estimating the values of τ and α are discussed in section 6.2.) The table reveals that the curves

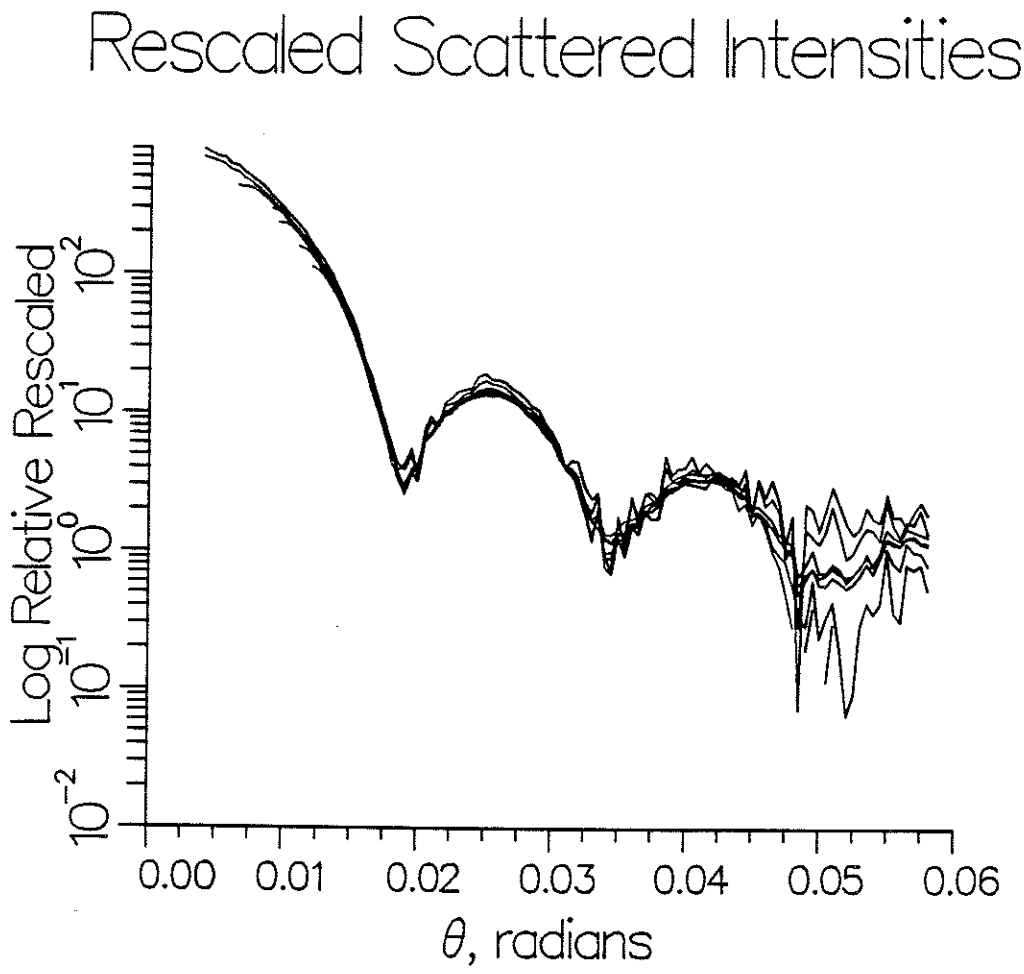

Figure 2.4: The measured, re-scaled light scattering from seven different concentrations of a Gaussian distributed size distribution. See text for further explanation.

Number of drops	Estimated Number Density	τ, Estimated	α, Estimated
1	154 ml^{-1}	0.01	.39
2	308 ml^{-1}	0.02	.79
4	617 ml^{-1}	0.04	1.6
6	925 ml^{-1}	0.06	2.4
8	1230 ml^{-1}	0.08	3.3
12	1850 ml^{-1}	0.13	5.0
16	2470 ml^{-1}	0.17	6.8

Table 2.1: Estimates of optical depth, τ and attenuation coefficient, α for verification of single scattering discussed in text.

in figure 2.4 represent high concentrations of particles. The estimates of attenuation, α provide an intuitive feel for how large these concentrations are. The attenuation in the table ranges from 0.39 to 6.8; contrast this with typical ocean waters, where α ranges from 0.03 for clear deep ocean water to 0.3 for the most turbid coastal water.

Furthermore, figure 2.4 and table 2.1 show that for these measurements, single scattering is valid for τ as large as 0.17, suggesting that the criterion $\tau \leq 0.1$ for single scattering is conservative. However, for smaller particles, which scatter relatively more into larger angles, the criterion might be less optimistic. Thus, in all of the experimental work, in addition to estimating τ when possible, I make a practice of overlaying plots of scattered intensities corresponding to different concentrations, with the same scale, to see that the curves follow each other. This insures that single scattering is a valid approximation for all the measurements.

2.5 Irregular, Randomly oriented Particles

In this work, I assume that small particles are spherical, at least in a statistical sense. Scattering can be found theoretically only for a few highly symmetrical shapes, e.g. for spheres, spherical shells [62], spheroids [78], or infinitely long cylinders. For all other shapes, one must solve Maxwell's equations numerically with appropriate boundary conditions, which is computationally expensive. Some investigators have considered modifying Mie theory for spheres to accommodate irregular, randomly oriented particles [79,80], but the resulting computational chore is at least as difficult as for spheres.

Although the scattering of individual particles varies with shape and orientation, an assembly of "equivalent" spheres can be found which has the nearly the same as scattering as an assembly of irregular particles. This is evident for the diffraction and Rayleigh approximations and, although not for large angle scattering [81], is supported by the work of a number of investigators [82,83] for Mie scattering.

The validity of assuming that a distribution of irregular particles has the same scattering as an equivalent distribution of spheres may depend on the validity of the diffraction approximation. Where the diffraction approximation is good, the assumption should be valid; otherwise, the validity of the assumption is a current research topic.

2.6 Effect of Index of Refraction

In later sections, I discuss the mathematical non-uniqueness of the inversion of $I(\theta)$ to yield $n(x)$. The models of scattering I use do not include the effect of index of refraction, as in the case of the diffraction approximation, or in the case of Mie scattering, do not assume a value for η. The mathematical and numerical uniqueness I discuss should be distinguished from that due to incomplete physical models of $I(\theta, x)$, i.e. the neglect of η.

To first order, at small θ, dependence on η is purely a multiplicative constant independent of θ, as illustrated in figures 2.5 and 2.6. This represents the most significant difference between Mie and diffraction theory at small θ, and is the basis of a number of improvements to diffraction theory. Fymat and Mease [76] noted that Shifrin and Punina proposed

$$I(\theta, x) = I_F(\theta, x)\,[Q_{ext}(x, \eta)/2]^2 \tag{2.37}$$

as an improvement, and presented their own refinement. $I_F(\theta, x)$ is the diffraction approximation scattered intensity and the multiplicative factor $Q_{ext}(x, \eta)$ is the extinction coefficient found from Mie theory. Fymat and Mease observed that for a distribution of particles of various refractive indices, the size distribution n^* found from inversion is actually a weighted integral over refractive index, i.e., that

$$n^*(x) = \int_0^\infty n(x)\beta(\eta)\,d\eta \tag{2.38}$$

where β is a probability distribution for η.

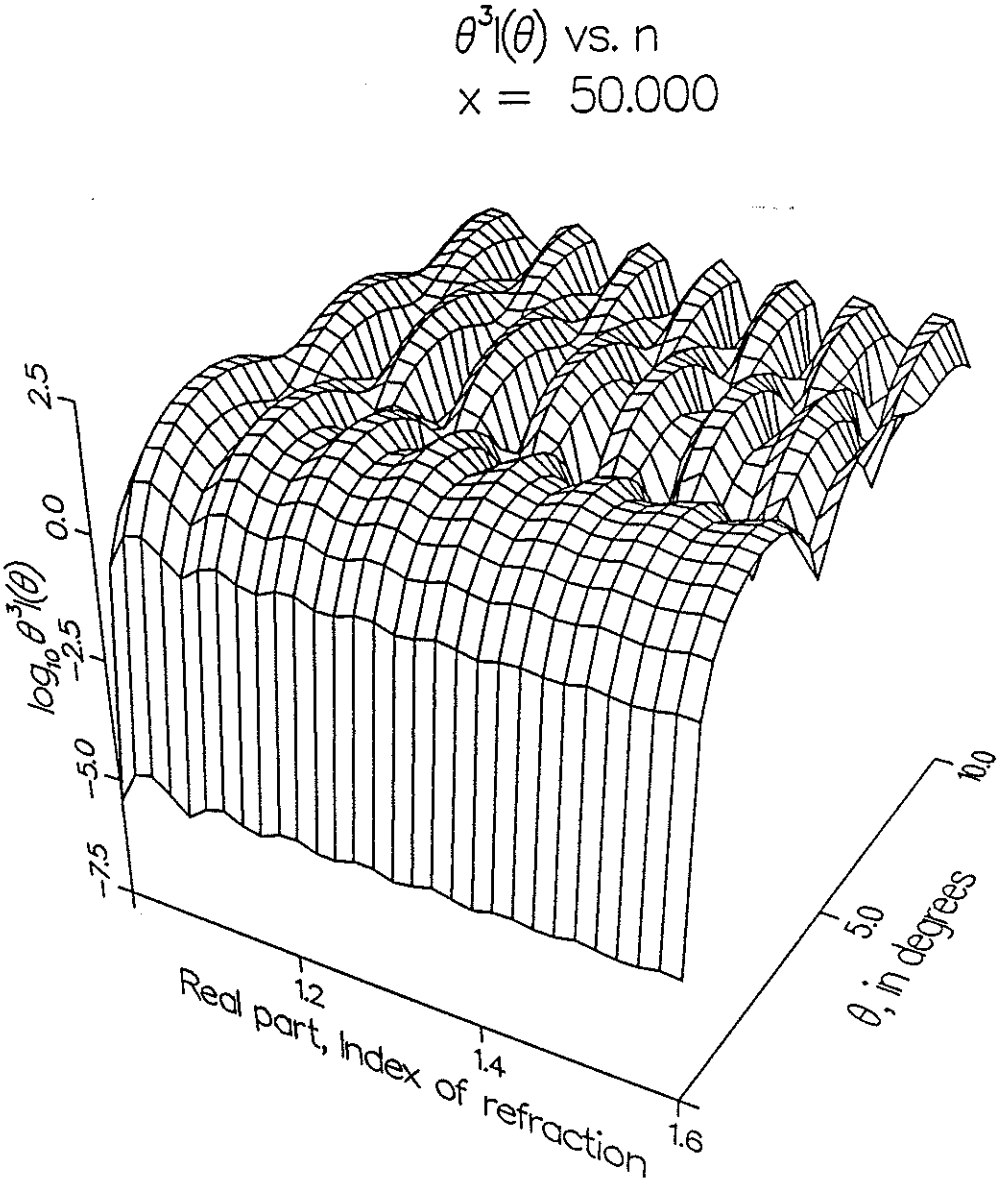

Figure 2.5: $\theta^3 I(\theta, x, \eta)$, for $x = 50$, $\eta'' = 0$, versus real part of index of refraction. Effect of η is almost purely multiplicative. $\theta^3 I(\theta, x, \eta)$ is plotted rather than $I(\theta, x, \eta)$ because it spans fewer orders of magnitude.

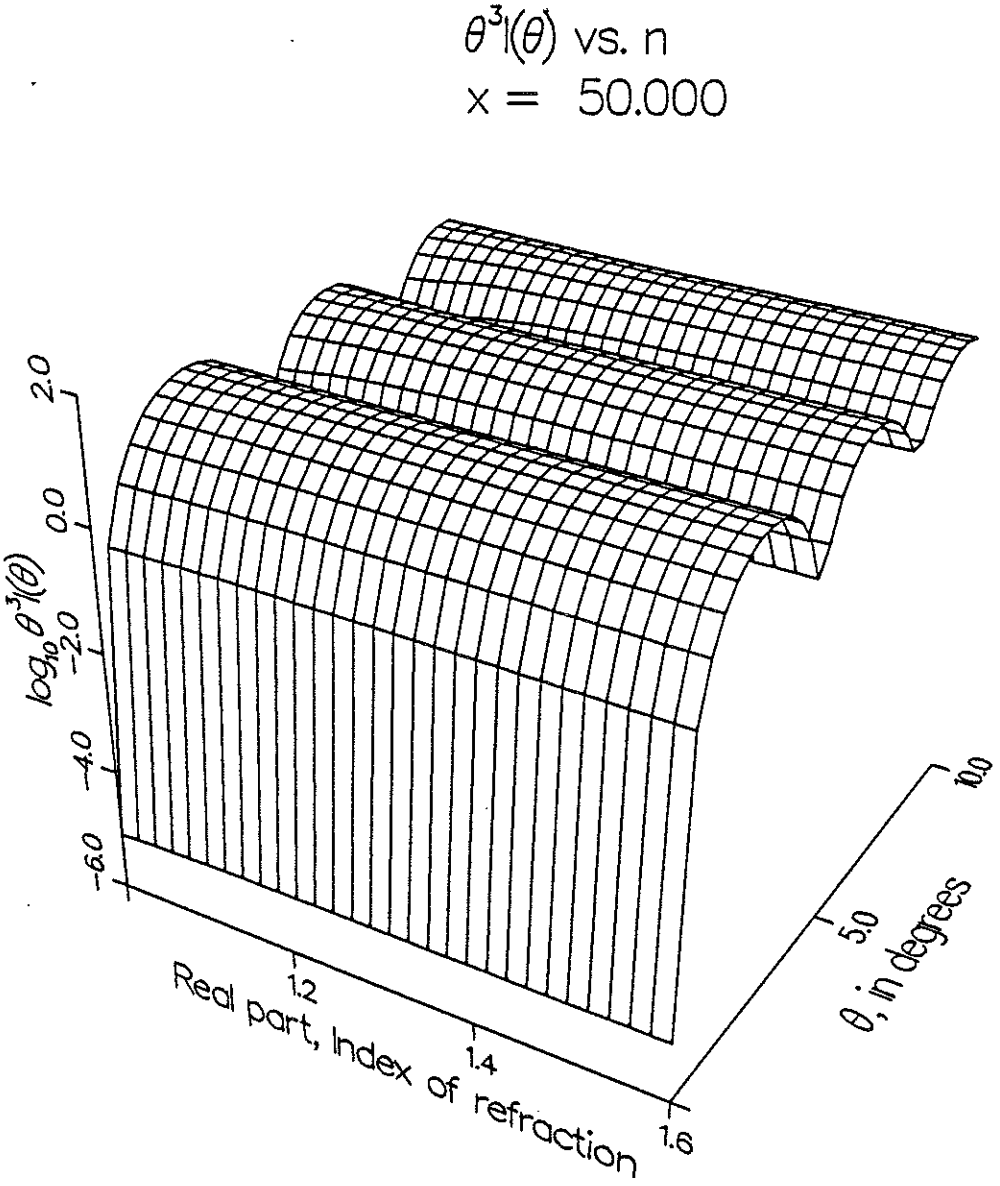

Figure 2.6: $I(\theta, x, \eta)$, for $x = 50$, $\eta'' = -0.1$, versus real part of index of refraction. Effect of η is negligible with this much attenuation.

2.7 Speckle Noise

Before proceeding, I note that a single realization of the intensity $I(\theta)$ is a sum of contributions from individual particles, which have randomly distributed phases. As such, the intensity is statistical in nature, and forms a "speckle" pattern. The statistical nature of $I(\theta)$ becomes apparent when the scalar field is written as a sum over contributions from individual particles:

$$u(\theta) = \sum_{i=1}^{N} u_i(\theta) = \sum_{i=1}^{N} |u_i(\theta)| e^{i\phi_i} \tag{2.39}$$

where there are N particles, with scalar amplitudes $|u_i(\theta)|$ and phases ϕ_i. The intensity is written in terms of the scalar field in the limit as time $T \to \infty$:

$$I(\theta) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} |u(\theta) e^{i\omega t}|^2 \, dt = |u_i(\theta)|^2 \tag{2.40}$$

From equation (2.39) one can derive the statistics of $I(\theta)$. Here I apply Goodman's approach for speckle in reflections from rough surfaces [84], to speckle in scattering from particles. The derivation starts with the following assumptions, which apply for a suspension of randomly located and oriented particles:

1. The $|u_i(\theta)|$ and ϕ_i of the $i-$th scatterer are statistically independent of each other and of the amplitudes and phases of all other particles.

2. The phases ϕ_i are uniformly and randomly distributed in the interval $(-\pi, \pi)$.

With these assumptions, the real and imaginary parts of the individual contributions $u_i(\theta)$ are statistically independent, and the sum in equa-

tion (2.39) becomes the classical random walk in the complex plane. Provided the number of particles N is large, the central limit theorem applies. As N becomes large, the real and imaginary parts $u^r(\theta)$ and $u^i(\theta)$ of $u(\theta)$ have Gaussian statistics, with joint probability density function:

$$p_{u^r,u^i}(u^r(\theta), u^i(\theta)) = \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{[u^r(\theta)]^2 + [u^i(\theta)]^2}{2\sigma^2}\right\} \tag{2.41}$$

where the variance σ^2 is

$$\sigma^2 = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1} N \frac{\langle |u_k(\theta)|^2 \rangle}{2} \tag{2.42}$$

The brackets $\rangle$ and $\langle$ denote the expectation operator.

The joint probability density function for $u^r(\theta)$ and $u^i(\theta)$ yields the joint probability density function of the intensity $I(\theta)$ and phase ϕ of the total field through a transformation. The real and imaginary parts of the scalar field are related to the intensity and phase by the transformation

$$u^r(\theta) = \sqrt{I(\theta)} \cos\phi \tag{2.43}$$

$$u^i(\theta) = \sqrt{I(\theta)} \sin\phi \tag{2.44}$$

The joint density function for intensity and phase is written in terms of these transformations and p_{u^r,u^i} as

$$p_{I,\phi}(I(\theta), \phi) = p_{u^r,u^i}(\sqrt{I(\theta)} \cos\phi, \sqrt{I(\theta)} \sin\phi) \|J\| \tag{2.45}$$

where $\|J\|$ is the Jacobian,

$$\|J\| = \left\| \begin{array}{cc} \frac{\partial u^r(\theta)}{\partial I(\theta)} & \frac{\partial u^r(\theta)}{\partial \phi} \\ \frac{\partial u^i(\theta)}{\partial I(\theta)} & \frac{\partial u^i(\theta)}{\partial \phi} \end{array} \right\| = \frac{1}{2} \tag{2.46}$$

and $\|\cdot\|$ represents absolute value. Performing the substitution in (2.45) yields

$$p_{I,\phi}(I,\phi) = \begin{cases} \frac{1}{4\pi\sigma^2}\exp\left(-\frac{1}{2\sigma^2}\right) & I \geq 0 \\ & -\pi \leq \phi \leq \pi \\ 0 & \text{otherwise} \end{cases} \tag{2.47}$$

Finally, the marginal probability density for the intensity is

$$p_I = \int_{-\pi}^{\pi} p_{I,\phi}\, d\phi = \begin{cases} \frac{1}{2\sigma^2}\exp\left(-\frac{I}{2\sigma^2}\right) & I \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.48}$$

The intensity $I(\theta)$ is thus described by negative exponential statistics. Such a distribution has variance equal to its mean, so that

$$p_I(I(\theta)) = \begin{cases} \frac{1}{2\langle I(\theta)\rangle^2}\exp\left(-\frac{I}{2\langle I(\theta)\rangle^2}\right) & I \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.49}$$

In summary, a single realization of the intensity is the ensemble average $\langle I(\theta)\rangle$ modulated by a random speckle component, $S(\theta)$:

$$I_s(\theta) = \langle I(\theta)\rangle S(\theta) \tag{2.50}$$

where $S(\theta)$ has negative exponential statistics of unit variance.

For a distribution of moving particles, as in hydrodynamics, intensity is averaged over a sufficient time interval that, assuming ergodicity, the observed intensity is $\langle I(\theta)\rangle$. In the experimental work, this is insured by stirring the samples being analyzed. Henceforth I drop the brackets $\rangle$ and $\langle$ denoting ensemble average. For stationary particles this is not valid, and requires the introduction of the speckle component as a source of noise [51,52].

2.8 Analysis of Scattering With a Thin Lens

In the introductory chapter, I stated that the lens in a diffraction particle sizing instrument serves as a Fourier transformer. The theory of section 2.2 provides the framework to show this. We will see that the intensity measured on the focal plane is the angular distribution of scattering. Readers familiar with Fourier optics will find that this is review material.

I will consider the action of a thin lens in terms of propagating a scalar field through a free space, through a thin lens, and through a space of length f, the focal length of the lens, onto a detector.

In section 2.2, we saw that the free space propagation of the scalar amplitude $u_0(x,y)$ through a distance z is given by the Fresnel diffraction integral

$$u(x,y,z) = \frac{i}{2\pi}\int_{-\infty}^{\infty} dx_0 \int_{-\infty}^{\infty} dy_0\, u_0(x_0,y_0) e^{-i\frac{k}{2z}[(x-x_0)^2+(y-y_0)^2]} \qquad (2.51)$$

For plane waves travelling through an aperture, the scalar amplitude $u_0(x,y)$ can represents the aperture function, that is,

$$u_0(x,y) = \begin{cases} 1 & \text{over the aperture} \\ 0 & \text{otherwise} \end{cases} \qquad (2.52)$$

We also saw that in the far-field, the Fraunhoffer approximation applies, and the field is the Fourier transform of the aperture function, $u_0(x,y)$.

The effect of a thin lens is to introduce a phase delay, that is, to slow the propagation of a wave. Thus the scalar amplitude u' emerging from a thin lens is

$$u'(x,y) = u(x,y) e^{i\frac{k}{2f}(x^2+y^2)} \qquad (2.53)$$

When used as a Fourier transformer, a thin lens is focused onto a detector. The field on the detector, resulting from a field represented by a scalar amplitude $u_0(x, y)$ at a distance z from the lens can be computed as follows:

1. The field is propagated through free space of depth z via the Fresnel integral, (2.51).

2. The field is propagated through the lens using equation (2.53).

3. The final propagation through free space of length f, the focal length of the lens, is performed using equation (2.51).

The algebra involved in this process simplifies somewhat, by recognizing that the Fresnel integral is a convolution, and can be treated by multiplying Fourier transforms using the Convolution theorem. The result is that the field on the detector is given simply by

$$u_d(x, y) = i\frac{2\pi k}{f} U_0\left(\frac{kx}{f}, \frac{ky}{f}\right) \tag{2.54}$$

where U_0 is the Fourier transform of $u_0(x, y)$.

When U_0 is circularly symmetric, that is,

$$U_0\left(\frac{kx}{f}, \frac{ky}{f}\right) = V_0\left(\frac{k}{f}\sqrt{x^2 + y^2}\right) = V_0\left(\frac{k}{f}s\right) \tag{2.55}$$

where s is defined as $s = \sqrt{x^2 + y^2}$, the field on the detector is also circularly symmetric, and can be written:

$$v_d(s) = i\frac{2\pi k}{f} V_0\left(\frac{ks}{f}\right). \tag{2.56}$$

Note that V_0 represents a *two-dimensional* Fourier transform, or *Hankel* transform. Furthermore, when V_0 represents an aperture function, its Fourier transform, v_d, shows up on the detection plane.

Thus, a lens produces a Fourier transform of an aperture, at its focal plane. We also saw, according to the Fraunhoffer approximation, that the scattering in the far field is also the Fourier transform, except for a constant phase term. The effect of a lens, is to produce the far field scattering of an aperture on the detection plane, with a scaling factor determined by the lens focal length.

Chapter 3

Theory: Asymptotic Analysis

3.1 Introduction

As we saw in chapter 2, the ensemble averaged scattered intensity $\langle I(\theta) \rangle$ from a suspension of single scattering particles can be written as an integral transform of the particle size distribution, $n(x)$:

$$\langle I(\theta) \rangle = \int_0^\infty I(\theta, x)\, n(x)\, dx \tag{3.1}$$

This is a Fredholm integral of the first kind, in which $x = kr$ is the particle radius r, non-dimensionalized by wave-number k, and θ is the scattering angle. The kernel $I(\theta, x)$ represents the scattering from a single particle. Henceforth I drop the brackets denoting ensemble average.

In the optical particle sizing problem, an inverse transform is sought so that $n(x)$ can be found from measurements of $I(\theta)$. The kernel $I(\theta, x)$ is usually modeled via Mie scattering theory for spheres, or by the diffraction approximation. The use of Mie theory requires that the particles be spherical in some statistical sense. In the diffraction approximation, $I(\theta)$ is

modeled as the Fraunhoffer diffraction from circular apertures of the same radii as the particles:

$$\langle I(\theta)\rangle = \frac{1}{k^3\theta^2}\int_0^\infty J_1^2(x\theta)\, x^2 n(x)\, dx \tag{3.2}$$

where J_1 is a first order Bessel function of the first kind.

Chin and Sliepcevich [8,9] derived an analytical transform to (3.2), which Fymat explored in further depth [10]:

$$n(x) = \frac{-2\pi k^3}{x^2}\int_0^\infty J_1(x\theta)Y_1(x\theta)x\theta\,\frac{d}{d\theta}\left[\theta^3 I(\theta)\right]\, d\theta \tag{3.3}$$

Here Y_1 is a first order Bessel function of the second kind. The infinite integration limit is meaningless physically, but, for many $n(x)$, the integrand falls off so rapidly in θ that one can substitute a reasonable limit with no effect. Since this transform is not essential to this analysis, I leave a derivation to appendix A.

In an experimental setting, $I(\theta)$ is sampled at discrete points θ_i. Unless $I(\theta)$ is smooth, the derivative of $\theta^3 I(\theta)$ makes numerical evaluation of (3.3) unstable. Also, the accuracy of the inversion depends on the validity of the diffraction approximation.

Many investigators have pursued purely numerical rather than analytical inverses to (3.1). Their work is based on rewriting (3.1) as the discrete sum

$$\langle I(\theta_i)\rangle = \sum_{j=1}^{N} n(x_j)\, I(\theta_i, x_j) W_{ij} \tag{3.4}$$

and in turn as the matrix equation:

$$\mathbf{I} = \mathbf{T}\mathbf{n} \tag{3.5}$$

where $\mathbf{I}_i = \langle I(\theta_i) \rangle$, $\mathbf{n}_j = n(x_j)$, $\mathbf{T}_{ij} = I(\theta_i, x_j) W_{ij}$ and $i = 1, \ldots, M$, and $j = 1, \ldots, N$. W_{ij} is a matrix representing the quadrature scheme; I choose simply $W_{ij} = \Delta x$.

At first glance, $\mathbf{n}$ could be found from

$$\mathbf{n} = \mathbf{T}^{-1}\mathbf{I}, \tag{3.6}$$

but the matrix $\mathbf{T}$ either lacks full rank, so that $\mathbf{T}^{-1}$ does not exist, or is nearly singular, so that $\mathbf{n}$ is unduly affected by noise in $\mathbf{I}$. A number of investigators have developed algorithms which avoid these problems by utilizing *a priori* information or otherwise imposing constraints on the solution $\mathbf{n}$. These algorithms include a number of least squares methods which include singular value decomposition, the Philips-Twomey algorithm, the Backus-Gilbert inversion technique, and a number of iteration schemes [85,86,87,88].

The purely numerical methods have the advantage that the scattering kernel $I(\theta, x)$ can be based on arbitrary scattering calculations, such as Mie scattering for spheres, which are more exact than the diffraction approximation. Furthermore, they avoid taking derivatives numerically. In addition, they can, in principle, tolerate missing data.

Rather than turn immediately to search for the "best" algorithm that finds $\mathbf{n}$, I return to equations (3.2) and (3.3), seeking intuitive understanding of the ill-posedness of the inverse transform, considering the issues of uniqueness and stability, sampling requirements, resolution versus noise, and algorithms for inverting $I(\theta)$.

In this chapter, I show that the functions $\theta^3 I(\theta)$ and $xn(x)$ are asymp-

totically related by a Fourier-sine transform, which is more stable than the transform relating $I(\theta)$ and $n(x)$. By exploiting the asymptotic relation to a Fourier transform, I derive a Nyquist criterion, based on the largest size in a distribution and the sampling interval in θ. I construct a matrix equivalent to T of equation (3.6) that is nearly non-singular, yet complete physically, and allows for efficient computation of $I(\theta)$ from $n(x)$ and vice-versa. These results, based on the diffraction approximation, are approximately valid for the scattering via other models as well, e.g. Mie scattering from spheres.

3.2 Asymptotic Analysis of the Diffraction Approximation

This section begins with an asymptotic analysis of (3.2), the diffraction approximation relating the particle size distribution $n(x)$ to the intensity $I(\theta)$. This analysis reveals the sources of uniqueness and instability, in the sense of Tikhonov [89], observed in inverting (3.2). We will see that to seek $xn(x)$ from $\theta^3 I(\theta)$ is equivalent to seeking $n(x)$ from $I(\theta)$, but is less susceptible to numerical instability. Observing that, in an asymptotic sense, $xn(x)$ and $\theta^3 I(\theta)$ are related by a Fourier-sine transform, I derive important properties using signal processing theory, e.g. a Nyquist sampling theorem. The analysis also motivates the choice of inversion algorithms. Finally, the theory sets the agenda for the experimental part of this work.

3.2.1 Uniqueness and Stability

By non-uniqueness, one means that there exists a non-trivial $\phi(x)$ such that

$$\int_0^\infty I(\theta, x)\phi(x)\, dx = 0 \tag{3.7}$$

If a non-trivial $\phi(x)$ exists, then an infinite number of solutions to (3.1) can be formed from the superposition $n(x) + \alpha\phi(x)$, where α is an arbitrary constant.

Mathematical instability means that small changes in $I(\theta)$, require large changes in the transform, $n(x)$. Mathematically, this can be stated in terms of the norms

$$l_I(I) = \left\{\int_0^\infty |I(\theta)|^2\, d\theta\right\}^{1/2} \tag{3.8}$$

and

$$l_n(n) = \max|n(x)| \text{ for } 0 \leq x \leq \infty \tag{3.9}$$

An inversion is unstable if, as the norm between two intensities I_1 and I_2 becomes small:

$$l_I(I_1 - I_2) \to 0,$$

the norm between the solutions n_1 and n_2 remains bounded:

$$l_n(n_1 - n_2) \geq L,$$

where L is a lower bound, $L > 0$. This is often stated that the linear transformation of $I(\theta)$ to $n(x)$ is not continuous. The generalization of continuity of functions to continuity of transformations of functions, and the conditions on a linear transformation such that a continuous (i.e. stable) inverse exist are discussed precisely by Naylor and Sell [90].

For our purposes, a loose, but practical definition of instability serves. Simply, a transformation is unstable if it is sensitive to small changes, or noise, in the function being transformed.

Asymptotic Analysis for Small Particles

In the analysis of (3.2), I first examine the contribution to $I(\theta)$ from the smallest particles. Of course, (3.2) is not appropriate for $x \ll 1$; the Rayleigh approximation is. Nonetheless, I perform the analysis using (3.2), because it is useful for numerical simulations, so a comparison to exact theory is useful. To do this, I break up (3.2) into two parts:

$$I(\theta) = \frac{1}{k^3\theta^2}\int_0^{x_{ref}} J_1^2(x\theta)\, x^2 n(x)\, dx + \frac{1}{k^3\theta^2}\int_{x_{ref}}^{\infty} J_1^2(x\theta)\, x^2 n(x)\, dx \qquad (3.10)$$

In the limit that $x_{ref}\theta \to 0$,

$$J_1^2(x_{ref}\theta) \sim x_{ref}^2\theta^2/4. \qquad (3.11)$$

Thus, the first integral in (3.10), representing the contribution from the smallest particles, becomes:

$$I_{x_{ref}\ll 1/\theta} = \frac{1}{4k^3}\int_0^{x_{ref}} x^4 n(x)\, dx \qquad (3.12)$$

This integral contributes little to $I(\theta)$ because $n(x)$ is weighted by x^4. As $x_{ref} \to 0$, it contributes nothing. Thus, $n(x)$ is not unique at $x = 0$ because $n(0)$ can take on any value with out changing $I(\theta)$. Moreover, for $0 < x \ll 1$, a large change in $n(x)$ results in an infinitesimal change in $I(\theta)$. So, in addition to the non-uniqueness for $x = 0$, $n(x)$ is unstably represented by $I(\theta)$.

I emphasize that the asymptotic analysis of (3.2) for $x_{ref}\theta \ll 1$ represents the scattering from real particles approximately, and accurately only for large spheres with indices of refraction close to that of the surrounding medium. For $x \ll 1$, the Rayleigh approximation is appropriate:

$$I(\theta) = \frac{k^4|\gamma|^2}{d^2}\cos^2\theta \tag{3.13}$$

where $I(\theta)$ represents the scattered light when the incident light (of unit intensity) is polarized parallel to the plane of scattering.[1] $|\gamma|$ is a function of index of refraction, shape, and radius, x [63]. For such particles, the angular dependence of $I(\theta, x)$ is separable from the size dependence, and can be written:

$$I(\theta, x) = f(\theta)g(x) \tag{3.14}$$

For a polydispersion, $I(\theta)$ becomes

$$I(\theta) = f(\theta)\int_0^\infty n(x)g(x)\,dx \tag{3.15}$$

The only information $I(\theta)$ contains regarding $n(x)$ concerns one of its integral moments. For spheres of index of refraction m,

$$\gamma = x^3\left(\frac{m^2-1}{m^2+1}\right) \tag{3.16}$$

so

$$I_{x_{ref}\ll 1/\theta} = \frac{k^4\cos^2\theta}{r^2}\left(\frac{m^2-1}{m^2+1}\right)^2\int_0^{x_{ref}} x^6 n(x)\,dx \tag{3.17}$$

[1] When the incident light is polarized perpendicular to the plane of scattering, the intensity is $I(\theta) = k^4|\gamma|^2/d^2$. Thus the arguments that follow remain the same except that there is no angular dependence $\cos^2\theta$.

For $x \ll 1$, then, $I(\theta)$ conveys information regarding only the sixth moment of $n(x)$. Our conclusions regarding instability and non-uniqueness hold then, though the instability is even worse than for the diffraction approximation.

Asymptotic Analysis for Large Particles

To perform the asymptotic analysis for large particles, I assume that $n(x)$ has compact support, i.e., that $n(x) = 0$ for $x \geq x_u$, where x_u is a bound on the largest size in the distribution. This assumption is reasonable for realistic particle size distributions. As $x_u\theta \to \infty$,

$$J_1^2(x_u\theta) \sim \frac{1}{\pi x\theta}(1 - \sin 2x_u\theta). \tag{3.18}$$

This is a good approximation for $x_u\theta > 4$; see fig. 3.1. Thus for $\theta \gg 1/x_u$,

$$I(\theta) = \frac{1}{\pi\theta^3 k^3}\left[\int_0^\infty xn(x)\,dx - \int_0^\infty \sin 2x\theta\, xn(x)\,dx\right] \tag{3.19}$$

This is valid for smaller θ than indicated, because the contribution to $I(\theta)$ for small x is very small. Approximately, then, $I(\theta)$ is composed of two terms; a constant term,

$$C = \int_0^\infty xn(x)dx, \tag{3.20}$$

and a Fourier-sine transform term,

$$\int_0^\infty xn(x)\,\sin 2x\theta\,dx. \tag{3.21}$$

The sine transform term is discussed in the next sections. If the constant term can be subtracted out, the inversion of $I(\theta)$ that yields $xn(x)$

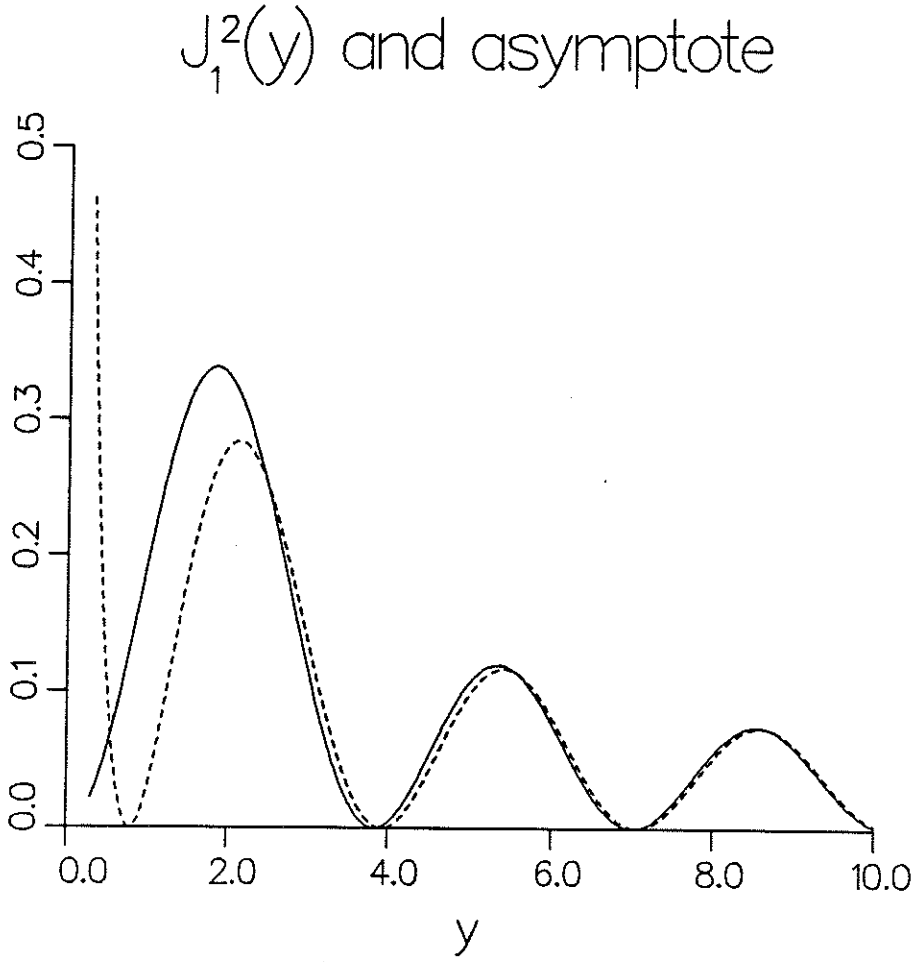

Figure 3.1: J_1^2 (solid line) compared with its asymptote (dashed line).

is mathematically stable, because the Fourier-sine transform is mathematically stable. The inversion that yields $n(x)$ is not stable however, because of instability as x becomes small (but remains large enough for (3.19) to remain valid). There is non-uniqueness in the limit that $x = 0$, but this is not an appropriate limit of (3.19). From (3.19), $\theta^3 I(\theta)$ and $xn(x)$ are asymptotically related by a sine transform, except for a constant additive term. Unlike the transform relating $I(\theta)$ and $n(x)$, which may be unstable for small sizes even in the limit $x_u\theta \gg 1$, this transform is well conditioned.

Rephrasing the Problem

I suggest a rephrasing of the optical particle sizing problem, so that $xn(x)$ is sought from $\theta^3 I(\theta)$ rather than $n(x)$ from $I(\theta)$. This has several advantages, listed in order of importance:

- Mathematical instability is *removed* in an asymptotic sense, in that there is stability for $x_u\theta \gg 1$. (i.e., for large particles).

- Mathematical instability is *reduced* in an asymptotic sense, in that for $x_{ref}\theta \ll 1$, $\theta^3 I(\theta)$ has stronger dependence on $xn(x)$ than on $n(x)$.

- $\theta^3\, I(\theta, x)/x$ has constant order of magnitude over θ and x, and furthermore, we will see that $\theta^3 I(\theta)$ varies less in order of magnitude over θ than does $I(\theta)$. This reduces dramatically the dynamic range requirements for computation, increasing numerical stability substantially.[2]

[2]This is particularly useful on the DEC VAX 11 series, for which FORTRAN-77 represents numbers from $-0.29\text{x}10^{-38}$ to $1.7\text{x}10^{38}$, for both double *and* single precision. Numbers less than $0.29\text{x}10^{-38}$ are rounded to zero. In my experience rounding to zero

The third advantage is the most important; one could seek $x^6 n(x)$ and eliminate mathematical instability altogether at small sizes, but would encounter numerical instability in the form of insufficient dynamic range.

Regrettably, an inversion of $\theta^3 I(\theta)$ loses the information contained in $I(0)$. In practice measurements at $\theta = 0$ can be impossible because of the large magnitude of the unscattered beam, so this may not be a problem. The next step may be to seek $xn(x)$ from a function $[\theta^3 + \rho(\theta)]I(\theta)$ rather than from $\theta^3 I(\theta)$. ρ might be the second term in an asymptotic series or an arbitrary but judiciously chosen function. In the meantime, I avoid $\theta = 0$.

The effect of seeking $n(x)$ from $\theta^3 I(\theta)$ is that some instability is removed to a separate, subsequent problem of finding $n(x)$ from $xn(x)$, or if one pleases, to the more stable problems of finding particle area and volume distributions $\pi x[xn(x)]$, and $4/3\pi x^2[xn(x)]$, respectively. One can think of a transform yielding $n(x)$ as one which demands from $I(\theta)$ information regarding particles of arbitrarily small size, which it does not have. No information is lost by demanding $xn(x)$ rather than $n(x)$.

3.2.2 Similarity to Sine Transform

An additional advantage of determining $\theta^3 I(\theta)$ from $xn(x)$ is that, except for the constant additive factor, this pair is approximately related by the well understood Fourier-sine transform. This became evident in equation (3.19). An asymptotic analysis of the inverse to (3.2) due to Chin and

seems to be the worst source of instability in inverting matrices — and this is not abetted by switching to double precision.

Sliepcevich [8,9] provides a similar result. For $y = x\theta \gg 1$, I substitute

$$J_1(y)Y_1(y)y = \frac{1}{\pi}\cos 2y + D(y), \tag{3.22}$$

where

$$D(y) = \frac{-3}{4\pi y}\sin 2y + \mathcal{O}(1/y^2) = \mathcal{O}(1/y) \text{as } y \to \infty, \tag{3.23}$$

into (3.3), finding

$$n(x) = \frac{-2k^3}{x^2}\int_0^\infty \cos 2x\theta \frac{d}{d\theta}\left[\theta^3 I(\theta)\right]\, d\theta + \frac{-2\pi k^3}{x^2}\int_0^\infty D(x\theta)\frac{d}{d\theta}\left[\theta^3 I(\theta)\right]\, d\theta \tag{3.24}$$

Figure 3.2.2 compares $J_1(y)Y_1(y)y$ with two of its asymptotes, $\frac{1}{\pi}\cos 2y$ and $\frac{1}{\pi}\cos 2y - \frac{3}{4\pi y}\sin 2y$. For $y > 3$, $\frac{1}{\pi}\cos 2y$ is an excellent approximation to $J_1(y)Y_1(y)y$, and $D(y) \approx 0$. For large x, (3.24) becomes:

$$n(x) = \frac{-2k^3}{x^2}\int_0^\infty \cos 2x\theta \frac{d}{d\theta}\left[\theta^3 I(\theta)\right]\, d\theta \tag{3.25}$$

Asymptotically, $x^2 n(x)$ and $\frac{d}{d\theta}[\theta^3 I(\theta)]$ are related by a Fourier-sine transform. This relation is easily derived from (3.19); the constant term in θ is eliminated by taking the derivative $\frac{d}{d\theta}[\theta^3 I(\theta)]$ and observing that it is asymptotically a cosine transform which is readily inverted, yielding (3.25). Equation (3.25) can be integrated by parts to eliminate the derivative in θ, yielding

$$n(x) = -\frac{2k^3}{x^2}\sin 2x\theta\, \theta^3 I(\theta)\Big|_{\theta=0}^{\infty} + \frac{4k^3}{x}\int_0^\infty \sin 2x\theta\, \theta^3 I(\theta)\, d\theta \tag{3.26}$$

Not surprisingly, except for the constant first term, $xn(x)$ is related to $\theta^3 I(\theta)$ by a Fourier-Sine transform, except for a constant additive term. It is not clear how to evaluate the first term at the infinite limit since $\sin 2x\theta$ is periodic, but this is not important in this development.

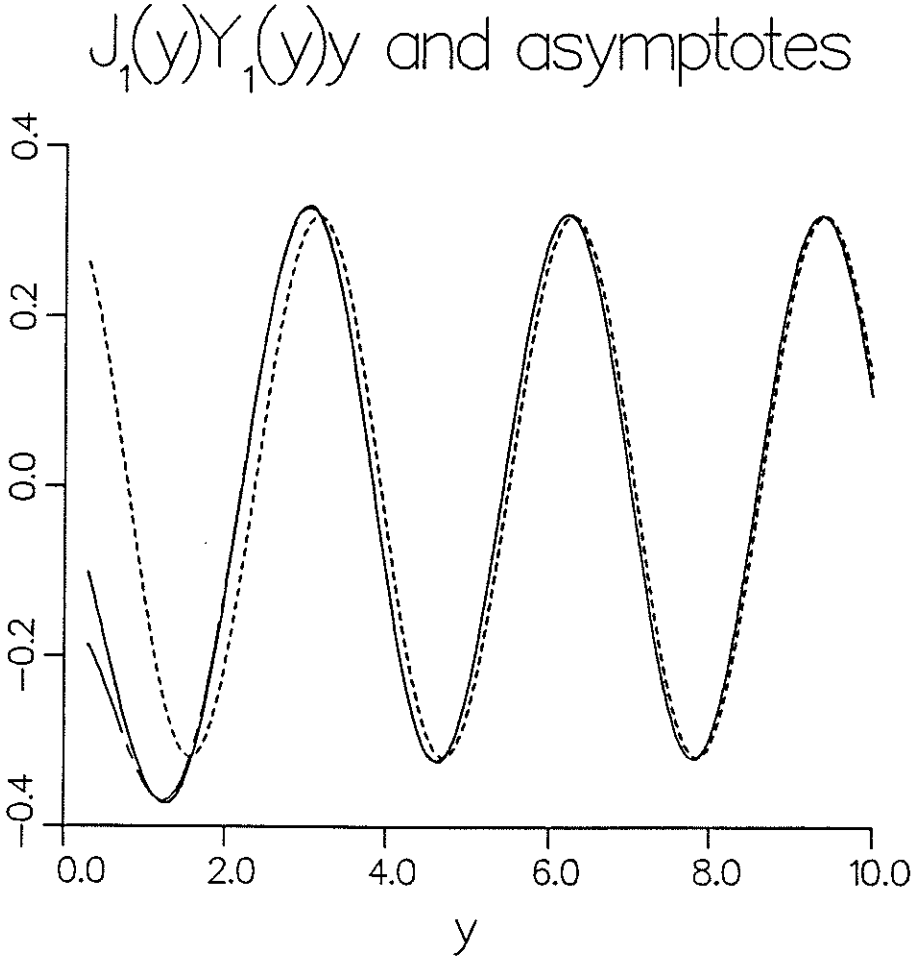

Figure 3.2: $J_1(y)Y_1(y)y$ compared with two asymptotes, $\frac{1}{\pi}\cos 2y$ and $\frac{1}{\pi}\cos 2y - \frac{3}{4\pi y}\sin 2y$.

3.2.3 Mathematical Precursor

For convenience, a change in notation and mathematical precursor are in order. I will regard $\theta^3 I(\theta)$ as data and denote it as $d(\theta)$; $d(\theta) = \theta^3 I(\theta)$. The result of any inversion schemes is a model, denoted $m(x) = xn(x)$. The complete set of functions $m(x)$ occupy a Hilbert space, equipped with the inner product

$$\langle u, v \rangle = \int_0^\infty u(x)v(x)\,dx, \tag{3.27}$$

which defines a norm,

$$l_2(u) = \langle u, u \rangle^{1/2}, \tag{3.28}$$

and transforms $m(x)$ to $d(\theta)$:

$$d(\theta) = \langle G(\theta, x), m(x) \rangle \tag{3.29}$$

where $G(\theta, x) = \theta^3 I(\theta, x)/x$ is the scattering kernel for the new notation. I assume that $m(x)$ has compact support, i.e. that $m(x) = 0$ for $x \geq x_u$. Asymptotically, for $x_u \gg 1$,

$$G(\theta, x) \approx \frac{1}{\pi k^3}(1 - \sin 2x\theta). \tag{3.30}$$

Thus, in the new notation, (3.19) becomes

$$d(\theta) \approx \frac{1}{\pi k^3}\int_0^{x_u} m(x)\,dx - \frac{1}{\pi k^3}\int_0^{x_u} m(x)\,\sin 2x\theta\,dx \tag{3.31}$$

3.2.4 Properties of the Relation Between $m(x) = xn(x)$ and $d(\theta) = \theta^3 I(\theta)$

The observation that $m(x)$ and $d(\theta)$ are, according to an asymptotic analysis of the diffraction approximation, related by a Fourier sine transform,

allows the application of elementary signal processing concepts. In this section I exploit the Fourier transform relation between $xn(x)$ and $\theta^3 I(\theta)$.

Band-limitedness of $d(\theta)$ and $I(\theta)$

Since $m(x)$ has compact support, the Fourier transform of $d(\theta)$ also has compact support; $d(\theta)$ is band-limited. I define band-limitedness in terms of the Fourier transform rather than in terms of equation (3.29):

Definition 1 *$d(\theta)$ is band-limited if its Fourier-transform,*

$$f(\omega) = \mathcal{F}\{d(\theta)\} = \int_{-\infty}^{\infty} d(\theta)\, e^{i\omega\theta}\, d\theta$$

has compact support.

A Fourier transform of $G(\theta, x) \approx \frac{1}{\pi k^3}(1 - \sin 2x\theta)$, reveals that $d(\theta)$ is approximately band-limited to $|\omega| \leq 2x_u$ by this definition. (Similarly, $I(\theta)$ is band-limited to $|\omega| \leq 2x_u$, due to the well known property of Fourier transforms, that if $F(\omega)$ is the Fourier transform of $I(\theta)$, then $\frac{d}{d\omega} f(\omega)$ is the transform of $i\theta I(\theta)$, where $i = \sqrt{-1}$.)

Of course, $d(\theta)$ has meaning for $0 < \theta < \infty$ *only* in the diffraction approximation. For real scattering, $I(\theta)$ is 2π-periodic in θ. $d(\theta)$ has meaning only at small angles. I denote the largest angle at which the diffraction approximation is valid by θ_{val}.

Though I have demonstrated that $d(\theta)$ is band-limited in the diffraction approximation, one can also demonstrate it for exact scattering from spheres as computed from Mie theory. In figure 3.3, I present $d(\theta) = \theta^3 I(\theta)$ for a sphere of size $x = 40$, and complex index of refraction $\eta = 1.125$ for

two polarizations from Mie scattering computations, and from the diffraction approximation. The region of validity of the diffraction approximation is evident. Figure 3.4 shows the frequency content of $d(\theta)$ in figure 3.3 as computed by the fast Fourier transform. Though there are substantial differences between the curves, they are similarly band-limited. Keeping in mind that the Fourier transform of $d(\theta)$ is asymptotically $m(x)$, except for a constant term, I have labeled the frequency axis approximately in terms of particle size x. The peaks of the curves all occur quite close to $x = 40$.[3] Some of the energy above $x = 40$ can be attributed to sampling of $I(\theta)$ over a subset rather than the full interval from 0 to π.

The band-limitedness of $I(\theta)$ and equivalently $d(\theta)$ can also be ascertained from Mie theory semi-mathematically. The scattered intensities of Mie theory are the functions $|S_1(\theta)|^2$ and $|S_2(\theta)|^2$. These can each be written in the form [64]:

$$S_{1 \text{ or } 2} = \sum_{n=1}^{\infty} \alpha_n \pi_n(\cos\theta) + \beta_n \tau_n(\cos\theta) \tag{3.32}$$

where α_n and β_n are coefficients that depend on x and η, and π_n and τ_n are defined in terms of Legendre Polynomials as follows:

$$\pi_n(\cos\theta) = \frac{1}{\sin\theta} P_n^1(\cos\theta) \tag{3.33}$$

$$\tau_n(\cos\theta) = \frac{d}{d\theta} P_n^1(\cos\theta) \tag{3.34}$$

[3] In fact, the figure suggests an algorithm for sizing single particles from measurements of $I(\theta)$. The size of a single particle can be determined by locating the peak in the FFT of $d(\theta)$.

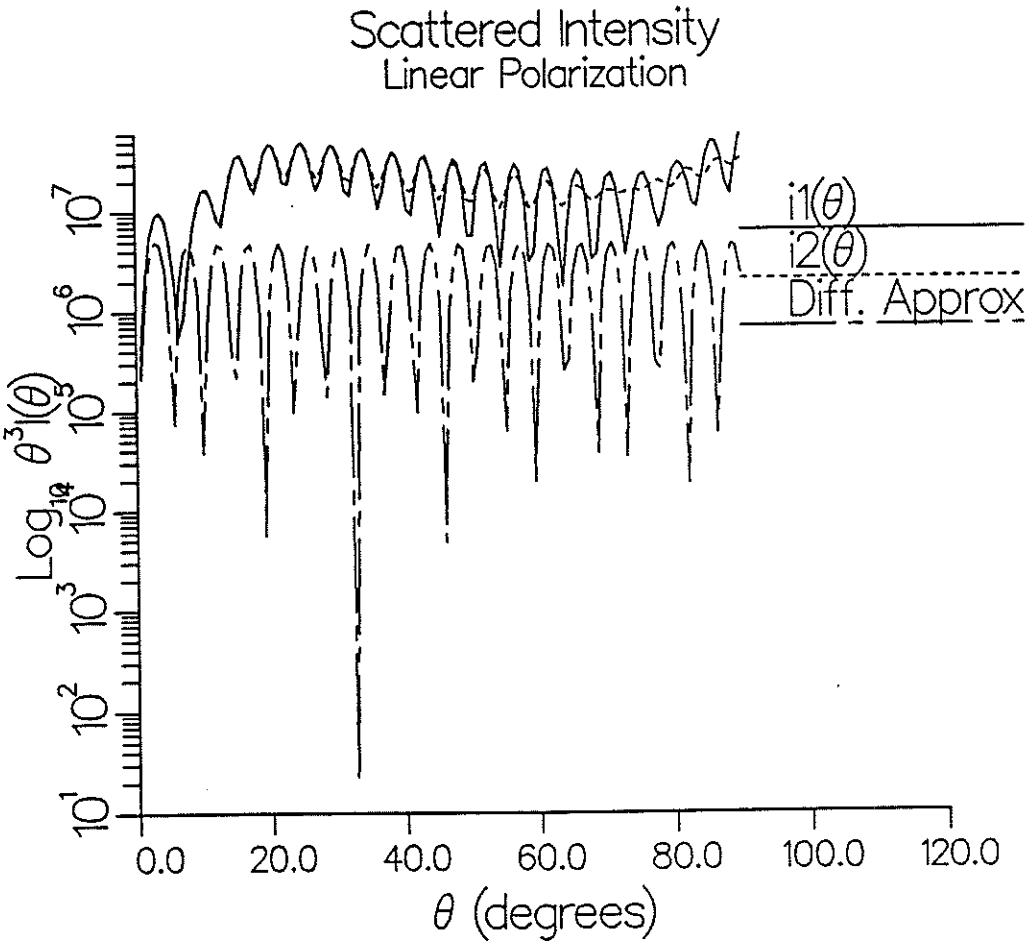

Figure 3.3: $d(\theta) = \theta^3 I(\theta)$ for a sphere with $x = 40$, and $\eta = 1.125$, for 2 polarizations from Mie scattering computations, and from the diffraction approximation. Of course, $d(\theta)$ has meaning only where the diffraction approximation is valid.

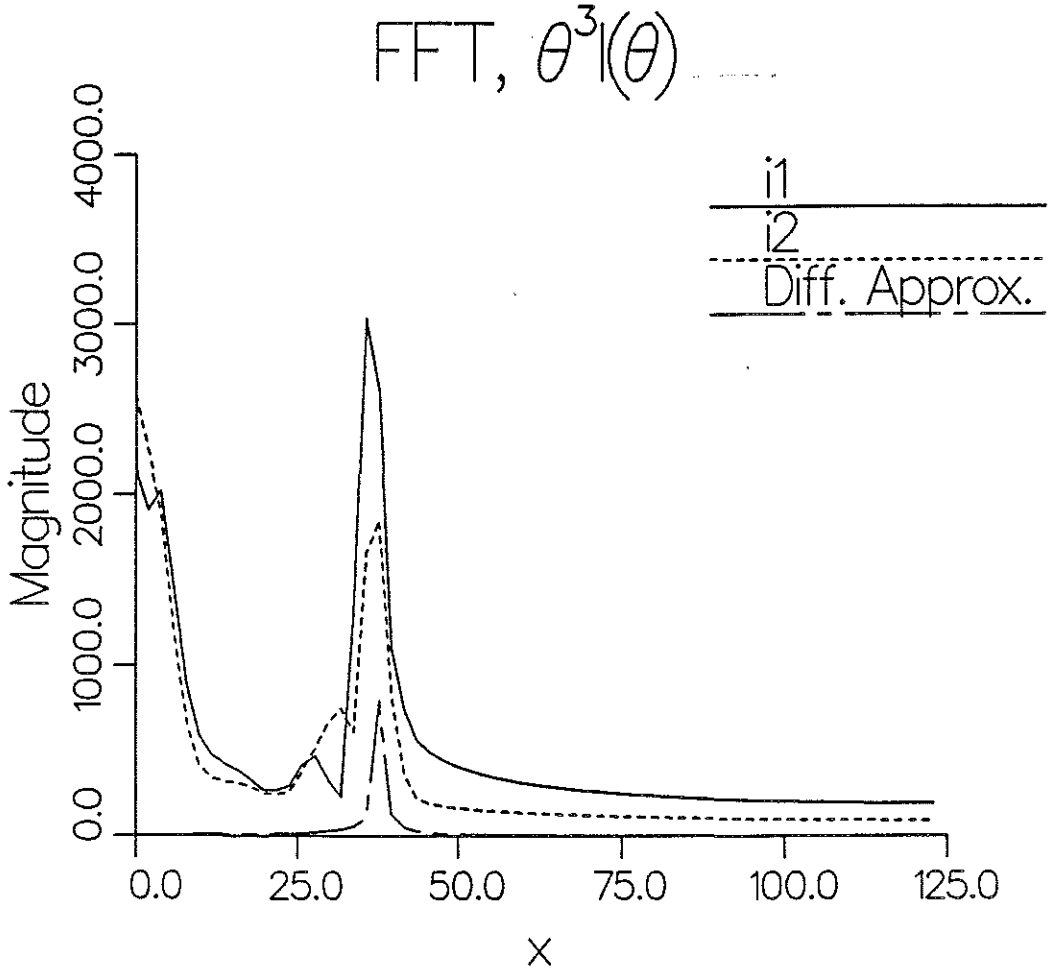

Figure 3.4: Fourier Transform of $d(\theta)$ of figure 3.3. The various scattering theories are similarly band-limited.

It can be shown that these can be written as

$$\pi_n(\cos\theta) = \sum_{n=1}^{N-1} q_n \cos n\theta \tag{3.35}$$

and

$$\tau_n(\cos\theta) = \sum_{n=1}^{N-2} q'_n \cos n\theta \tag{3.36}$$

where the q_n and q'_n are constant coefficients. Empirically one finds that the sum in equation (3.32) can be terminated at $n \approx x$, by computing the α_n and β_n for a number of particle sizes. Approximately, then,

$$|S_{1\ or\ 2}|^2 = \left| \sum_{n=1}^{n \approx x} p_n \cos n\theta \right|^2 \tag{3.37}$$

Where again the p_n are coefficients. This can be rewritten in a form

$$|S_{1\ or\ 2}|^2 = \sum_{n=1}^{n \approx x} r_n \cos 2n\theta \tag{3.38}$$

which approximately has $\cos 2x\theta$ as its highest frequency component.

In general, the approximate band-limitedness $|\Omega| \leq 2x_u$ is likely to be slightly optimistic for the diffraction theory from which it is derived, and somewhat more optimistic for real scattering, as predicted by Mie scattering. Though band-limitedness was derived for small θ in the diffraction approximation, the semi-mathematical treatment for Mie theory suggests that it may be approximately valid for larger angles as well.

Approximate Compact Support of $d(\theta)$

I assume that $m(x)$ is square integrable (i.e., that $\langle m, m \rangle$ is bounded), to show that $d(\theta)$ can have compact support, which will prove useful in developing a sampling theorem for $d(\theta)$.

For large θ the second term of (3.31) becomes small relative to the first, according to the Riemann–Lebesgue lemma [91], provided that $m(x) = xn(x)$ is square integrable. The rate at which it vanishes depends on $m(x)$, i.e., whether it is continuous, merely piecewise continuous, or has singularities, for example.[4] At large θ, $d(\theta)$ is asymptotic to a constant $C = \int_0^\infty m(x)\ dx$. (Equivalently, $I(\theta) \sim \theta^{-3}C$).

Suppose that, at θ greater than some angle θ_u, dependent on experimental precision or expediency, $d(\theta)$ is indistinguishable from the constant C. In an experiment, additional measurements at $\theta > \theta_u$ may provide little information regarding $m(x)$ other than refining the estimate of its average. Measurements at these angles specify $m(x)$ uniquely only to the extent that they constrain $d(\theta)$ to approach a constant.

In this case $d(\theta)$, and equivalently $I(\theta)$ need not be sampled for $\theta > \theta_u$, because it can be reconstructed trivially from $d(\theta_u)$, from $d(\theta) = d(\theta_u)$ for $\theta > \theta_u$. In terms of experimental work this is very useful because θ_u can be quite small; this is a justification to sample only at near-forward angles. A generalized definition of compact support for $d(\theta)$, is useful:

Definition 2 *$d(\theta)$ has* generalized *compact support if, for $\theta > \theta_u$, $d(\theta)$ is indistinguishable (within the precision of measurement) from a constant C to which it is asymptotic.*

Those who prefer the traditional definition of compact support may equivalently use the compact support of the derivative of $d(\theta)$ in the arguments

[4] $m(x)$ may not appear smooth for sparse distributions. For sparse distributions, I assume that a smooth $m(x)$ may be obtained by averaging a number of realizations of $d(\theta)$.

that follow, with no loss of validity.

An approximate lower bound on θ_u may be determined from (3.31) and the Heisenberg uncertainty principle. It is well known that if a function $f(\theta)$ has variance σ_θ^2 and its Fourier transform $F(x)$ has variance σ_x^2, the product of the variances has a lower bound:

$$\sigma_\theta \sigma_x \geq \frac{1}{2} \tag{3.39}$$

Equality holds if $f(\theta)$ and $F(x)$ are Gaussian. By substituting a Gaussian $m(x)$ into the second term of equation (3.31), the lower bound is found:

$$\theta_u \geq \frac{1}{4x_u} \tag{3.40}$$

I caution that equation (3.31) (and (3.19)) is not valid for $\theta > \theta_{val}$, the limit of validity of the diffraction approximation. (Recall that the diffraction approximation is based on the Fraunhoffer approximation, which is valid only for small θ.) Fortunately, an angle large enough for (3.31) to be valid can still be small enough for the diffraction approximation to be valid. That is, $d(\theta)$ can become indistinguishable from a constant within the region of validity of the diffraction approximation; $\theta_u < \theta_{val}$. I present two examples:

- The Gaussian distribution:

$$xn(x) = e^{\frac{(x-x_0)^2}{2\sigma^2}},$$

  where σ is small so that $xn(x)$ represents a "broad" Gaussian. For such a distribution, the second term in (3.19) is a narrow Gaussian,

with width inversely proportional to σ^2, and with sinusoidal modulation of frequency dependent on x_0. Thus, θ_u depends on σ^2, and will fall within the region of the diffraction approximation if σ^2 is large enough.

- $n(x) = \delta(x - x_o)$, where δ is the Dirac delta function. In this case, the second term in (3.19) is periodic. $I(\theta)$ *never* approaches an asymptote with in the region of validity of the diffraction approximation.

If $d(\theta)$ never reaches an asymptote within the region of validity of the diffraction approximation, that is, if $\theta_u > \theta_{val}$, then the region of support for $d(\theta)$ may have to be regarded as lying between $\theta = 0$ and $\theta = \pi$.

I present $\theta^3 I(\theta)$ for some physically unlikely but nonetheless instructive $n(x)$ in figures 3.5, 3.6 and 3.7. These figures demonstrate that the angle θ_u, at which $\theta^3 I(\theta)$ or $I(\theta)$ are indistinguishable from their asymptotes depends on the width of the size distribution and whether or not it is continuous.

3.2.5 Essential Sampling Interval of $d(\theta)$ and $I(\theta)$

The band-limitedness of $d(\theta)$ indicates that it can, in principle, be reconstructed uniquely from a discrete sampling of $d(\theta)$ at angles θ_i. The approximate generalized compact support of $d(\theta)$ determines the sampling interval within which $d(\theta)$ should be measured.

A consequence of the band-limitedness of $d(\theta)$, and of its approximate generalized compact support, is that $d(\theta)$ can be represented approximately by a finite number M of its samples, at angles $\theta_i > 0$, $i = 1, \ldots, M$. The number of samples, M, will be the subject of this section. The accuracy of

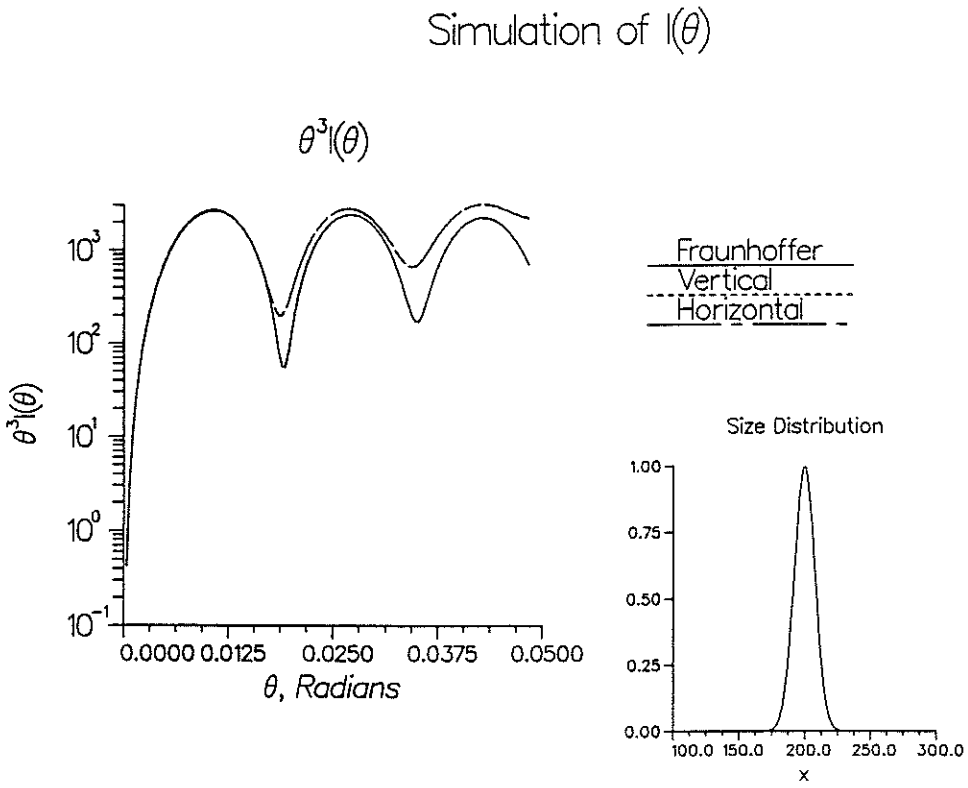

Figure 3.5: $\theta^3 I(\theta)$ for $n(x)$ a Gaussian, with $x = 200$, and $\sigma = 50$. θ_u falls well within the region of validity of the diffraction approximation. The legend "vertical" refers to $i_1(\theta)$, "horizontal" refers to $i_2(\theta)$, and "Fraunhoffer" refers to the diffraction approximation. The size distribution is shown in the inset plot. For the Mie scattering computations, $\eta = 1.125 - i0.01$.

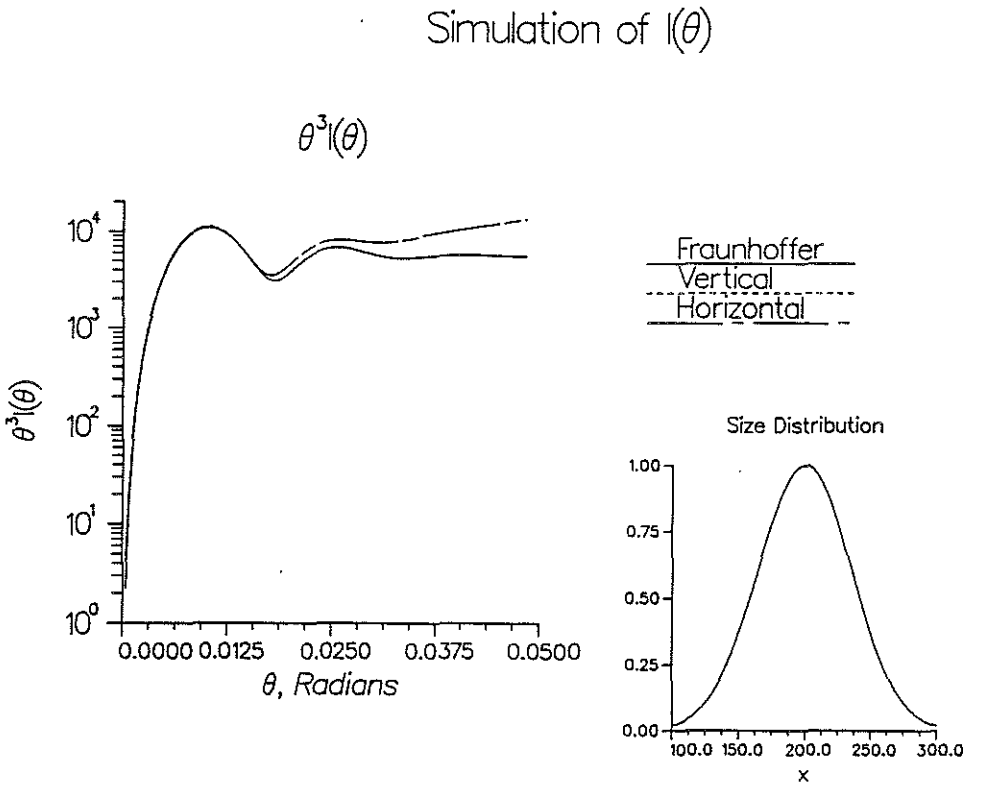

Figure 3.6: $\theta^3 I(\theta)$ for $n(x)$ a Gaussian, with $x = 200$ and $\sigma = 11$. θ_u is larger than for $\sigma = 50$. The size distribution for this plot is similar to one for which I have measured and inverted scattering experimentally. The angles θ nearly correspond to those of the experimental measurements. Also, see the caption to figure 3.5 for further explanation.

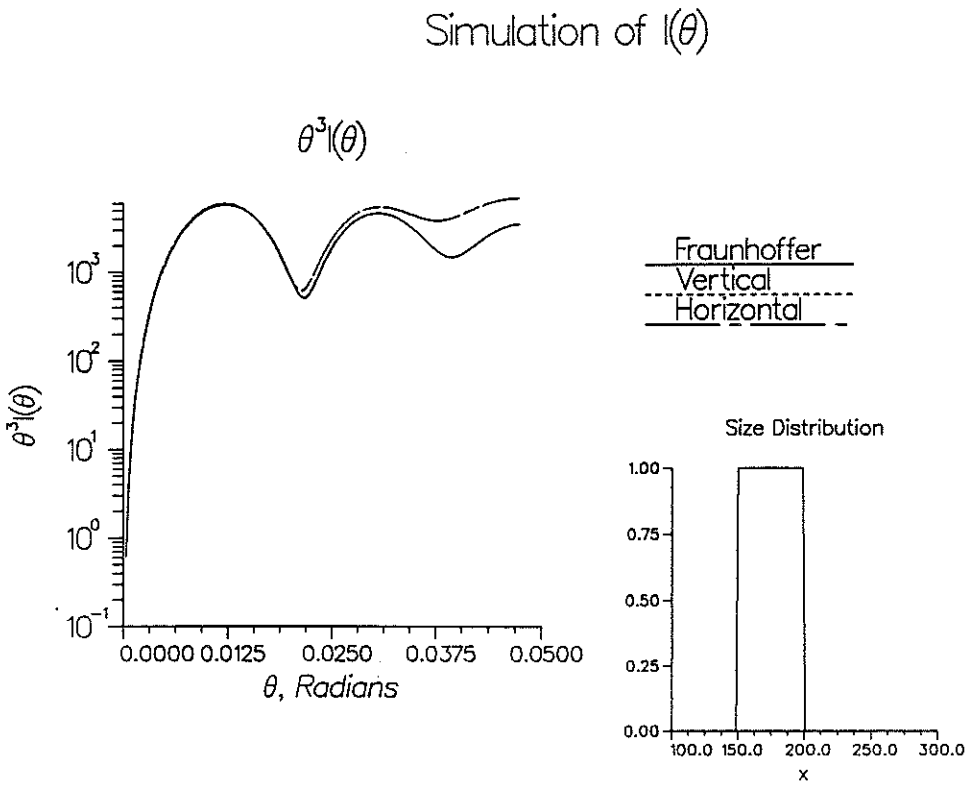

Figure 3.7: $\theta^3 I(\theta)$ for $n(x) = 1$ for $150 \leq x \leq 200, n(x) = 0$ otherwise. A discontinuous size distribution results in an intensity which approaches its asymptote more slowly than a continuous distribution. See the caption to figure 3.5 for further explanation.

this approximation depends on the approximation of compact support of $d(\theta)$.

Because $m(x)$ is a unique transform of $d(\theta)$ for $x > 0$, $d(\theta)$ and $m(x)$ are equivalent representations. Hence, sampling of $m(x)$ at M points $x_j > 0$ uniquely represents $d(\theta)$. That is, if $d(\theta)$ can be reconstructed from M points of itself, it can also be reconstructed from M points of $m(x)$.

There is approximate duality between $d(\theta)$ and $m(x)$, which results from the approximate Fourier relation between the two functions. By duality, one means that anything that is true about a function $f(x)$ from the properties of its Fourier transform $F(\omega)$ is also true about $F(\omega)$ from the properties of $f(x)$. For example, if $f(x)$ is band-limited $F(\omega)$ has compact support; by duality, if $f(x)$ has compact support, $F(\omega)$ is band-limited. Because of the approximate duality of $m(x)$ and $d(\theta)$, I conjecture that $m(x)$ can be uniquely reconstructed from M samples of $d(\theta)$, provided that both $d(\theta)$ and $m(x)$ have compact support.

Minimum Sampling Rate of $d(\theta)$ (or $I(\theta)$)

For robust inversion or reconstruction of $d(\theta)$ from a finite number of samples, $d(\theta)$ should be sampled evenly in θ over the region of support of $d(\theta)$, excluding $\theta = 0$. The region of support of $d(\theta)$ may be $0 < \theta < \pi$ if $d(\theta)$ does not have compact support according to definition 2. In this section I present a Nyquist theorem for sampling $d(\theta)$, or equivalently $I(\theta)$, based on x_u and θ_u.

I assume that $d(\theta)$ has compact support, i.e, according to definition 2, $d(\theta) = C$ for $\theta_{valid} > \theta > \theta_u$. Otherwise one can set $\theta_u = \pi$ and the

discussion that follows will still be valid for real-life scattering, in which $I(\theta)$ is even and 2π-periodic.

From the band-limitedness of $d(\theta)$, one can state the following:

Statement 1 *$d(\theta)$ is uniquely determined if uniformly sampled in the interval $0 < \theta \leq \theta_u$ at sample points with separation*

$$\Delta\theta \leq \frac{\pi}{2x_u}. \tag{3.41}$$

By *uniquely determined* I mean that $d(\theta)$ can be reconstructed unambiguously at all θ. Equivalently,

Statement 2 *$m(x)$ is uniquely determined from $d(\theta)$ for $x > 0$ if $d(\theta)$ is sampled in the interval $0 < \theta \leq \theta_u$ at sample points with separation*

$$\Delta\theta \leq \frac{\pi}{2x_u}. \tag{3.42}$$

Stated differently,

Statement 3 *Given a detector array that samples evenly from $\theta = 0$ to θ_u, not including $\theta = 0$, there is a* Nyquist-size *particle*

$$x_{Nyquist} = \frac{\pi}{2\Delta\theta} \tag{3.43}$$

such that the energy from larger particles is aliased.

In practice it is usually not possible to sample evenly from $\theta = 0$ to θ_u. For reasons I discuss in other sections, it is particularly difficult to sample near $\theta = 0$. Instead, it is common to sample from a minimum angle θ_{min} to a maximum θ_{max}. For sampling of $d(\theta)$ in an interval $\theta_{min} \leq \theta_i \leq \theta_{max}$, where

$\theta_{min} \geq 0$ and $\theta_{max} \leq \theta_u$, generalized sampling theorems are required. I refer the reader to the literature on generalizations of the Shannon sampling theorem for uneven sampling [92,93]. I make use of a generalized Shannon sampling theorem, stated as follows:

Statement 4 *If a finite number of uniformly distributed points of a band-limited function $f(\theta)$ are migrated to new positions θ_i', forming a new distribution, $f(\theta)$ remains uniquely defined and can be reconstructed, provided that the* average *sampling rate is twice the Nyquist Frequency.*

So, I will regard a sampling of $d(\theta)$ from θ_{min} to θ_{max} as being a sampling from $\theta = 0$ to θ_u, in which the samples are migrated to new positions. In principle then, $d(\theta)$ can be reconstructed from M of its samples if the average sample interval, $\overline{\Delta\theta}$:

$$\overline{\Delta\theta} = \frac{\theta_u}{M-1} \tag{3.44}$$

samples according to the criterion:

$$\overline{\Delta\theta} \leq \frac{\pi}{2x_u} \tag{3.45}$$

One can generalize the sampling theorems as follows:

Statement 5 *$d(\theta)$ is uniquely determined if uniformly sampled in the interval $0 < \theta_{min} \leq \theta \leq \theta_{max} \leq \theta_u$ at sample points with separation*

$$\Delta\theta \leq \frac{\pi(M-1)}{2x_u} \tag{3.46}$$

where M is the number of samples in θ. Equivalently,

Statement 6 *$m(x)$ is uniquely determined from $d(\theta)$ for $x > 0$ if $d(\theta)$ is sampled in the interval $\theta_{min} \leq \theta \leq \theta_{max}$ at sample points with separation*

$$\Delta\theta \leq \frac{\pi(M-1)}{2x_u} \tag{3.47}$$

Stated differently,

Statement 7 *Given a detector array that obtains M samples evenly from $\theta = \theta_{min}$ to θ_{max}, there is a* Nyquist-size *particle*

$$x_{Nyquist} = \frac{\pi(M-1)}{2\theta_u} \tag{3.48}$$

such that the energy from larger particles is aliased.

Some sample calculations show that compact support of $d(\theta)$ allows the extraction of considerably more information from $d(\theta)$ than the limited knowledge that $I(\theta)$ is even and 2π-periodic. Consider a sampling of $d(\theta)$ at 128 points each separated by 0.1°. If $\theta_u = \pi$, $x_{Nyquist} = 64$ or $r_{Nyquist} = 6.4\,\mu m$ for $\lambda = 632.8$ nm. These are very small particles. On the other hand, if the array is known to sample up to θ_u, i.e. if $\theta_u = 12.8°$, $x_{Nyquist} = 900$ or $r_{Nyquist} = 90\,\mu m$.

There is some partial, empirical verification of the Nyquist criterion in the literature, for Mie scattering, as opposed to the diffraction approximation from which the criterion is derived. J. V. Dave [86] performed inversions on data obtained by Setzer, who measured the scattering from a polydispersion of spheres over large angles, actually, from 12° to 146° at 1° intervals. For $\Delta\theta = 1°$, $\theta_u = 124°$, our criterion gives $x_{Nyquist} = 90$, or $r_{Nyquist} = 9\,\mu m$. Dave [94] performed his inversion for $I(\theta, x)$ computed via Mie Scattering for spheres, and concluded that

> Knowledge of the upper limit of the size distribution [is important] especially when the distribution has a sharp cutoff.

He found this cutoff to be $7\,\mu m$, which is close to my estimate of $x_{Nyquist} = 9\,\mu m$, derived from the diffraction approximation which is valid only at small angles.

3.2.6 Stability of Inverting Unevenly Spaced Data

Uneven sampling of $d(\theta)$ is of interest experimentally. In an experiment, the unscattered beam is orders of magnitude stronger than the scattered beam at very small angles, so it is not possible to sample at these angles. Furthermore, a detector array may be too short to sample $d(\theta)$ up to θ_u.

We saw in the previous section that $d(\theta)$ can be reconstructed from M, or inverted to obtain $m(x)$, if the average sample interval, $\overline{\Delta\theta}$ obeys a sampling criterion. The issue of stability of inverting unevenly spaced data remains, however.

In practice, a band-limited function $d(\theta)$ is reconstructed from its samples at θ_i with greatest immunity to noise if the samples are evenly spaced over the entire interval of the function. The reconstruction of a function $d(\theta)$ from uneven samples is unstable, and is a current research topic, otherwise known as *extrapolation of band-limited signals from unevenly spaced samples.*

The dual research topic, *spectral analysis with unevenly spaced data*, is also an unstable problem, and is of interest to us. Since $m(x)$ is so closely related to the Fourier transform of $d(\theta)$, the inversion yielding $m(x)$ from

M uneven samples of $d(\theta)$ is unstable as well.

In addition to inverting unevenly spaced samples of $d(\theta)$, I consider extrapolating (or interpolating) $d(\theta)$ into θ where it is not sampled. Such extrapolation (interpolation) could be a way to reconstruct the Volume Scattering Function (VSF) at small θ from measurements at larger angles. Moreover, an understanding of the instability of band-limited extrapolation provides understanding of the instability of inverting $d(\theta)$ from unevenly spaced samples.

Let us consider the stability of reconstructing or inverting $d(\theta)$ with missing or unevenly spaced data. I consider the related, simpler problems, of band-limited extrapolation of $d(\theta)$, and determining the Fourier transform, $D(\omega)$. Note that $m(x)$ is closely related to $D(\omega)$. This can be seen by substituting the definition of the Fourier transform into the approximate Fourier-sine relation between $m(x)$ and $d(\theta)$, to yield

$$m(x) \approx K/x - k^3 i[D(2x) - D^*(2x)]. \tag{3.49}$$

Here $i = \sqrt{-1}$, '$*$' denotes complex conjugate, and K is a constant. If determining $D(\omega)$ is unstable, then so is inverting for $m(x)$, because the two are essentially equivalent. This allows us to draw upon a wealth of literature, including the paper by Slepian and Pollak [95].

I assume $d(\theta)$ of diffraction theory to be square integrable over the interval $-\infty < \theta < \infty$, so that it possesses a Fourier transform. Furthermore, assume that $d(\theta)$ is band-limited, i.e. $D(\omega) = 0$ for $|\omega| \geq \Omega$. Suppose that

one has a piece of $d(\theta)$:

$$\tilde{d}(\theta) = \begin{cases} d(\theta) & \text{for } |\theta| \leq \theta_{max} \\ 0 & \text{otherwise} \end{cases}$$

We want to determine $d(\theta)$ for $|\theta| > \theta_{max}$ from $\tilde{d}(\theta)$. I consider this simpler case for simplicity, rather than the general case in which, in addition, $\tilde{d}(\theta) = 0$ for $|\theta| \leq \theta_{min}$.

Slepian and Pollak showed that there are functions ψ_i, $i = 1, \ldots, \infty$ for any $\theta_{max} > 0$, and any Ω, which have the following properties:

- The ψ_i are orthonormal, band-limited to Ω, and complete in the space of square integrable, band-limited functions:

$$\int_{-\infty}^{\infty} \psi_i(\theta)\psi_j(\theta)\, d\theta = \delta_{ij} \tag{3.50}$$

  for $i, j = 0, \ldots, \infty$. δ_{ij} is the Kronecker delta function; $\delta_{ij} = 1$ for $i = j$, and $\delta_{ij} = 0$ for $i \neq j$.

- The ψ_i are orthogonal in the interval $-\theta_{max} \leq \theta \leq \theta_{max}$ in addition to the larger interval, and are complete in the space of functions square integrable from $-\theta_{max}$ to θ_{max}.

$$\int_{-\theta_{max}}^{\theta_{max}} \psi_i(\theta)\psi_j(\theta)\, d\theta = \lambda_i \delta_{ij} \tag{3.51}$$

- The ψ_i are solutions to

$$\lambda_i \psi_i(\theta) = \int_{-\theta_{max}}^{\theta_{max}} \frac{\sin \Omega(\theta - s)}{\pi(\theta - s)} \psi_i(s)\, ds \tag{3.52}$$

These remarkable functions are the *prolate spheroidal functions.* These and the eigenvalues λ_i are functions of $c = \Omega\theta_{max}$; i.e., $\lambda_i = \lambda_i(c)$, $\psi_i = \psi_i(c, \theta)$.

$d(\theta)$ can be expanded in terms of the ψ_i:

$$d(\theta) = \sum_{n=0}^{\infty} a_n \psi_n(\theta) \tag{3.53}$$

where

$$a_n = \int_{-\infty}^{\infty} d(\theta)\psi_n(\theta)\, d\theta \tag{3.54}$$

Due to the orthogonality of the ψ_i in the smaller interval, the a_n may also be found in terms of $\tilde{d}(\theta)$:

$$a_n = \frac{1}{\lambda_n} \int_{-\theta_{max}}^{\theta_{max}} \tilde{d}(\theta)\psi_n(\theta)\, d\theta \tag{3.55}$$

In principle, then, one can approximate $d(\theta)$ for all θ using

$$d_N(\theta) = \sum_{n=0}^{N} a_n \psi_n(\theta) \tag{3.56}$$

with mean square error

$$\int_{-\infty}^{\infty} [d(\theta) - d_N(\theta)]^2\, d\theta = \sum_{n=N+1}^{\infty} a_n^2 \tag{3.57}$$

The mean square error can be made arbitrarily small by making N large. The mean square error in $(-\theta_{max}, \theta_{max})$ is

$$\int_{-\theta_{max}}^{\theta_{max}} [d(\theta) - d_N(\theta)]^2\, d\theta = \sum_{n=N+1}^{\infty} a_n^2 \lambda_n \tag{3.58}$$

The eigenvalues λ_i diminish quite rapidly for $n > 2c/\pi$, as shown in table 3.1. Because

$$\int_{-\infty}^{\infty} \psi_i(\theta)\psi_i(\theta) = 1 \tag{3.59}$$

n	$c = 0.5$	$c = 1.0$	$c = 2.0$	$c = 4.0$	$c = 8.0$
0	$3.1\text{x}10^{-1}$	$5.7\text{x}10^{-1}$	$8.8\text{x}10^{-1}$	$1.0\text{x}10^{0}$	$1.0\text{x}10^{0}$
1	$8.6\text{x}10^{-3}$	$6.3\text{x}10^{-2}$	$3.6\text{x}10^{-1}$	$9.1\text{x}10^{-1}$	$1.0\text{x}10^{0}$
2	$3.9\text{x}10^{-5}$	$1.2\text{x}10^{-3}$	$3.6\text{x}10^{-2}$	$5.1\text{x}10^{-1}$	$1.0\text{x}10^{0}$
3	$7.2\text{x}10^{-8}$	$9.2\text{x}10^{-6}$	$1.2\text{x}10^{-3}$	$1.1\text{x}10^{-1}$	$9.6\text{x}10^{-1}$
4	$7.3\text{x}10^{-11}$	$3.7\text{x}10^{-8}$	$1.9\text{x}10^{-5}$	$8.8\text{x}10^{-3}$	$7.5\text{x}10^{-1}$
5	$4.6\text{x}10^{-14}$	$9.5\text{x}10^{-11}$	$1.9\text{x}10^{-7}$	$3.8\text{x}10^{-4}$	$3.2\text{x}10^{-1}$
6	$2.0\text{x}10^{-17}$	$1.7\text{x}10^{-13}$	$1.4\text{x}10^{-9}$	$1.1\text{x}10^{-5}$	$6.1\text{x}10^{-2}$
7	$6.6\text{x}10^{-21}$	$2.2\text{x}10^{-16}$	$7.0\text{x}10^{-12}$	$2.3\text{x}10^{-7}$	$6.1\text{x}10^{-3}$
8	$1.6\text{x}10^{-24}$	$2.1\text{x}10^{-19}$	$2.8\text{x}10^{-14}$	$3.6\text{x}10^{-9}$	$4.2\text{x}10^{-4}$

Table 3.1: Eigenvalues λ_i of $\psi_i(c,\theta)$. From Slepian *et al.*, 1961.

and

$$\int_{-\theta_{max}}^{\theta_{max}} \psi_i(\theta)\psi_i(\theta) = \lambda_i, \tag{3.60}$$

small eigenvalues imply that $\psi_i(\theta)$ has the most energy outside the interval $(-\theta_{max}, \theta_{max})$. Thus, while the error of equation (3.58) becomes small rapidly as $N \to \infty$, the error over the larger interval of (3.57) can still be large. Consequently, extrapolation of $d(\theta)$ in to angles where it is not sampled can be unstable.

One can also consider the problem of determining $D(\omega)$ from $\tilde{d}(\theta)$, which is similar to finding $m(x)$ from $\tilde{d}(\theta)$. Fourier transforming (3.53):

$$D(\omega) = \sum_{n=0}^{\infty} a_n \psi_n(\omega) \tag{3.61}$$

Applying Parseval's relation to (3.57),

$$\int_{-\infty}^{\infty} [D_N(\omega) - D(\omega)]^2 \, d\omega = \sum_{n=N+1}^{\infty} a_n^2 \tag{3.62}$$

As with (3.57), this can remain large as the error in the measured interval given by (3.58) becomes arbitrarily small. Thus, the Fourier transform

from limited data is unstable, even when unique. Since $m(x)$ is closely related to the Fourier transform, $D(\omega)$, through the relation $\mathrm{m(x)} \approx K/x - k^3 i[D(2x) - D^*(2x)]$, it follows that inversion for $m(x)$ is unstable as well.

A number of authors have proposed algorithms which reduce instability of band-limited extrapolation, by introducing additional constraints on the solution, such as positivity, for example [96,97,98,99,100]. Even with constraints the inversions tend to be sensitive to noise. If however, $d(\theta)$ is measured over a large enough portion of the interval from $\theta = 0$ to θ_u, the instability is surmountable.

3.3 Resolution

The approximate Fourier transform relation between $m(x)$ and $d(\theta)$ also determines the resolution and variance in any transformation from one to the other. By resolution I refer to the minimum separation Δx between which peaks in $m(x)$ are distinguishable. Variance refers to the noise sensitivity of a transformation of real data.

It is well known in time series analysis that the best resolution obtainable in a spectrum estimated from data of finite length is

$$\Delta f = \frac{1}{T} \tag{3.63}$$

where T is the length over which data is measured and Δf is the resolution in frequency [101]. Since this is the best resolution obtainable, it comes at the expense of sensitivity to noise. A least squares estimate with no damping yields such an estimate; an example would be an FFT of the data

with no windowing (that is, a rectangular window with length T)

Equation (3.63) can be applied to the particle sizing problem, using the approximate Fourier-sine relation between $m(x)$ and $d(\theta)$. The best resolution one can expect is thus

$$\Delta x = \frac{1}{2(\theta_{max} - \theta_{min})}, \tag{3.64}$$

provided, of course, that the Nyquist criterion is met.

With *a priori* information, such as a lower bound size x_{min} and and an upper bound x_{max}, it would seem that better resolution is possible, given simply by

$$\Delta x = \frac{(x_{max} - x_{min})}{M}, \tag{3.65}$$

where, M is the number is the number of samples in θ. In fact, this is not the case. If an algorithm exists which determines the size distribution to higher resolution than predicted by equation (3.64), then it implicitly extrapolates the scattered intensity into angles where it is not measured. The algorithm does this by providing a size distribution from which the scattering at unmeasured angles can be computed. Since such extrapolation is unstable, the higher resolution is unattainable.

The Backus–Gilbert formalism is commonly used in inverse theory for examining the tradeoff between resolution and variance. Since this formalism is also an inversion algorithm, I discuss it in section 4.2.2.

3.4 Finite Width Detectors.

I assume that intensity is available as point samples. Since real detectors have elements of finite width, this may not be the case. This is true for the instrument made by Malvern Instruments, Ltd. [4,5,6] Then, rather than measuring intensities at specific angles, one measures a set of integrals of intensity of the form

$$D_i = \int_{\theta_i - 1}^{\theta_i} d(\theta)\, d\theta \tag{3.66}$$

This operation is a convolution. The effect is the same as sampling a low-pass filtered intensity, and the validity of the preceding sections remains unchanged.

In my experimental work, I use EG&G Reticon photodiode arrays. For these devices, individual diodes are sufficiently small (25μm) that infinitesimal width of detectors is a good approximation.

Chapter 4

Numerical Transformations

In section 3.2.1, I rephrased the optical particle sizing problem, to seek $m(x) = xn(x)$ from $d(\theta) = \theta^3 I(\theta)$, which I argued are related by a more stable transform because, except for a term constant in θ, it is asymptotically a Fourier-sine transform. This similarity of the $m(x) \leftrightarrow d(\theta)$ transform pair to a Fourier transform can be exploited to form efficient, noise resistant algorithms for finding $m(x)$ from $d(\theta)$ and vice versa. An eigen-analysis allows us to investigate the conditioning of the $m(x) \leftrightarrow d(\theta)$ transform versus the $n(x) \leftrightarrow I(\theta)$ transform.

4.1 The Numerical Problem

An experiment yields $d(\theta)$ at M points θ_i. The data, $d_i = d(\theta_i)$, is related to the model $m(x)$, a continuous function, through the inner product

$$d_i = \langle g_i(x), m(x) \rangle \tag{4.1}$$

where the $g_i(x)$ are linearly independent and represent the scattering kernel; $g_i(x) = \theta_i^3 I(\theta_i, x)/x$. The inverse problem is to determine $m(x)$. This would seem to be an infinitely under-determined problem, because d_i is discrete and $m(x)$ is continuous. If, however, $m(x)$ has compact support and $d(\theta)$ is sufficiently sampled over its region of support, then, according to the approximate results of section 3.2.5, $m(x)$ is uniquely determined.

4.1.1 Continuous Estimates of $m(x)$

Whether or not $m(x)$ is uniquely determined by the samples d_i, an estimate of $m(x)$, which I denote $\hat{m}(x)$, can be constructed from a linear combination of the functions $g_i(x)$:

$$\hat{m}(x) = \sum_{i=1}^{M} \alpha_i g_i(x) \tag{4.2}$$

A unique set of coefficients α_i can be found by satisfying $d_i = \langle g_i, m \rangle$ and minimizing some norm of $\hat{m}(x)$. The norm $\langle \hat{m}(x), \hat{m}(x) \rangle$ in terms of the inner product of equation (3.27) is minimized when

$$\alpha = \mathbf{\Gamma}^{-1}\mathbf{d} \tag{4.3}$$

Substitution of (4.3) into (4.2) yields the well-known least squares result, $\hat{m}_{ls}(\theta)$. [102] Here α is the column M-vector formed from the α_i, $\mathbf{d}$ is formed from d_i, and $\mathbf{\Gamma}_{ij}$ is the correlation or "Gram" matrix, with the elements

$$\mathbf{\Gamma}_{ij} = \langle g_i, g_j \rangle \tag{4.4}$$

According to the sampling theorems of section 3.2.5, $m(x)$ is uniquely specified for $x > 0$ by the data d_i when $x_u < x_{Nyquist}$, where the Nyquist size

$x_{Nyquist}$ depends on the sampling of $d(\theta)$. When uniquely specified, $\hat{m}(x) = m(x)$ for $0 < x < x_{Nyquist}$. Otherwise, $\hat{m}(x)$ will depend on the norm that is minimized, or on other constraints that are imposed, in addition to its dependence on $d_i = \langle g_i, m \rangle$. An estimate $\hat{m}(x)$ could be found which has least integrated absolute value for example, or least maximum value, in addition to the least squares estimate $\hat{m}_{ls}(x)$. It is easy to derive least square estimates based on different norms; a common example is a norm based on the derivative of $m(x)$, so that $\hat{m}(x)$ is the smoothest function, in some sense, fitting the data.

Unless $m(x)$ is uniquely specified by (4.1), the estimates $\hat{m}(x)$ may resemble $m(x)$ but no two in general will be the same. All estimates $\hat{m}(x)$ are constrained in some way, but the constraints are not always appropriate. That an estimate be positive, for example, might be a more appropriate constraint than that it be smoothest in some sense, for example. Unfortunately, positivity is seldom a sufficient constraint. In general $\hat{m}(x)$ must be found that satisfies a constraint, minimum norm for example, whether or not it is appropriate.

4.1.2 Discrete Estimates of $m(x)$

Unfortunately, for the optical particle sizing problem, it is not possible to construct the Gram matrix Γ analytically because the integrations represented by the inner products $\langle g_i, g_j \rangle$ are intractable, for g_i from both diffraction and Mie theory. By discretizing $g_i(x)$ and $m(x)$ at N sample

points x_j and writing integrals as discrete sums, (4.1) becomes

$$\mathbf{d} = \mathbf{G}\mathbf{m}, \tag{4.5}$$

where $\mathbf{m}_j = \mathbf{m}(x_j)$ and $\mathbf{G}_{ij} = \theta_i^3/x_j I(\theta_i, x_j)\Delta x$, $i = 1, \ldots, M$, and $j = 1, \ldots, N$.

In this discrete formulation an estimate $\hat{\mathbf{m}}$ is sought of $\mathbf{m}$. So that the discrete sums of (4.5) are good approximations to the integrals in (4.1), the integration increment Δx is usually made as small as feasible. The result is that (4.5) represents an over-determined set of equations ($N > M$); $\mathbf{G}^{-1}$ does not exist.

Nonetheless a *pseudo-inverse* $\mathbf{H}$ can be found so that

$$\hat{\mathbf{m}} = \mathbf{H}\mathbf{d} \tag{4.6}$$

may give a good estimate of $\mathbf{m}$. For example, $\mathbf{H}$ can be found from the continuous least squares estimate of the previous section:

$$\mathbf{H}_{ls} = \mathbf{G}^{\mathsf{T}}\boldsymbol{\Gamma}^{-1} \tag{4.7}$$

The Gram matrix $\boldsymbol{\Gamma}$ can be determined approximately as $\boldsymbol{\Gamma} = \mathbf{G}\mathbf{G}^{\mathsf{T}}$:

$$\mathbf{H}_{ls} \approx \mathbf{G}^{\mathsf{T}}(\mathbf{G}\mathbf{G}^{\mathsf{T}})^{-1}, \tag{4.8}$$

which is the classic least squares inverse for under-determined discrete systems.

I leave a catalog of algorithms for $\mathbf{H}$ and estimates of $m(x)$ to another section. My approach here is to construct an non-singular (or nearly so) MxM matrix $\mathbf{Q}$ such that

$$\mathbf{d} = \mathbf{Q}\mathbf{m} \tag{4.9}$$

is an alternate approximation to (4.1). The justification for seeking such a matrix is that M points of either $d(\theta)$ or $m(x)$ are sufficient to uniquely specify both functions (provided $m(x)$ and $d(\theta)$ have compact support and are appropriately sampled), and are linearly related.

The matrix **Q** I construct is simply **G** with sample points x_j chosen judiciously. Recognizing the approximate role of **G** as — except for a constant additive factor — a Fourier-sine transformer, I choose the points x_j so that **G** is approximately the same matrix as that which performs a discrete sine transform. Next I justify the large integration increment Δx. This approach is similar to that of Hagin, [103], who numerically solved a Fredholm integral that is asymptotically a Fourier-cosine transform.

Once constructed as an MxM matrix, **G** will have two sources of singularity. The first is due to the mathematical non-uniqueness and instability at small sizes x, because $d(\theta)$ and **d** contain little or no information regarding the smallest particles. The second is aliasing; this is avoided if the x_j lie in the interval $(0, x_{Nyquist}]$.

In all further discussion, I assume that $d(\theta)$ has, to good approximation, compact support according to the (special) definition 2, and that it is sampled according to or more densely than the Nyquist criterion. Thus, $d(\theta)$ can be reconstructed to good approximation (depending on the assumption of compact support of $d(\theta)$) from either the M samples d_i or M samples of $m(x)$. Of course, I assume that $x_j \neq 0$, and $\theta_i \neq 0$.

Construction of G: Evenly Spaced Data θ_i.

First, I construct G for M samples of $d(\theta)$ at θ_i evenly spaced between 0 and θ_u, but not including $\theta = 0$. In practice, one would hope that the largest angle sampled θ_{max} would exceed the angle at which $d(\theta)$ is indistinguishable from its asymptote, and set $\theta_u = \theta_{max}$. If $\theta_{max} < \theta_u$, one simply accepts the resulting error. Approximately,

$$d(\theta) = \theta^3 I(\theta) \approx \frac{1}{\pi k^3} \int_0^\infty m(x)\ (1 - \sin 2x\theta)\ dx \tag{4.10}$$

From this, the elements of G are given approximately as

$$\mathrm{G}_{ij} \approx \frac{\Delta x}{\pi k^3}(1 - \sin 2x_j\theta_i) \tag{4.11}$$

For convenience, I define a matrix G′, to which the elements of G are asymptotic, except for a multiplicative factor:

$$\mathrm{G}'_{ij} = (1 - \sin 2x_j\theta_i) \tag{4.12}$$

According to the diffraction approximation, the elements of G'_{ij} are good approximations to $\frac{\pi k^3}{\Delta x}\mathrm{G}_{ij}$ for $\theta_i x_j > 3$.

Given θ_i from an experiment, I choose the x_j so that G′ is invertible. This does not guarantee that G, when calculated by Mie or diffraction theory will have an inverse, but I expect that it will be less singular than if the x_j's are chosen haphazardly.

Initially, it is not clear how to choose the x_j so that G′ is invertible. I write G′ as the difference of two matrices,

$$\mathrm{G}' = \mathrm{U} - \mathrm{V}, \tag{4.13}$$

where $\mathbf{U}_{ij} = 1$, and

$$\mathbf{V}_{ij} = \sin 2x_j\theta_i. \tag{4.14}$$

One can easily choose the x_j so that $\mathbf{V}$ has an inverse, by making it orthogonal, as I show later. When $\mathbf{V}$ is orthogonal, we will see that $\mathbf{G}'$ is invertible.

To make $\mathbf{V}$ orthogonal, I choose the x_j's so that $\mathbf{V}$ is the matrix representation of the discrete Fourier-sine transform. Starting with the well known Fourier transform pair for discrete sequences F'_k and f'_n,

$$F'_k = \sum_{n=0}^{N'-1} f'_n \, e^{-i2\pi kn/N'} \tag{4.15}$$

$$f'_n = \frac{1}{N'} \sum_{k=0}^{N'-1} F'_k \, e^{i2\pi kn/N'} \tag{4.16}$$

it is easy to derive a discrete sine transform pair for two odd sequences. In terms of (4.15) and (4.16), F'_k and f'_n are odd when $F'_k = -F'_{-k} = -F'_{N-k}$, and $f'_n = -f'_{-n} = -f'_{N-n}$, respectively. The appropriate transform pair is:

$$F_k = 2\sum_{n=1}^{M} f_n \sin \frac{2\pi kn}{N'} \tag{4.17}$$

$$f_n = \frac{2}{N'} \sum_{n=1}^{M} F_k \sin \frac{2\pi kn}{N'} \tag{4.18}$$

where $k = 1,\ldots,M$, $n = 1,\ldots,M$, and M is the number of points to be transformed. N' is related to M as follows:

$$N' = \begin{cases} 2M+2 & \text{for } M \text{ odd} \\ 2M+1 & \text{for } M \text{ even} \end{cases}$$

In matrix form, (4.17) and (4.16) become

$$\mathbf{F} = \mathbf{S}\mathbf{f} \tag{4.19}$$

$$\mathbf{f} = \mathbf{S}^{-1}\mathbf{F} \tag{4.20}$$

where

$$\mathbf{S}_{ij} = \sin\frac{2\pi ij}{N'} \tag{4.21}$$

$$\mathbf{S}_{ij}^{-1} = \frac{4}{N'}\sin\frac{2\pi ij}{N'} \tag{4.22}$$

with $i = 1,\ldots,M$ and $j = 1,\ldots,M$. The inverse $\mathbf{S}^{-1}$ is known to exist from (4.17) and (4.18), and from the fact that $\mathbf{S}$ is a well-known orthogonal matrix. (In fact, the matrix $\mathbf{R} = \sqrt{N'}\mathbf{S}$ is *orthonormal*, meaning that $\mathbf{R}^{-1} = \mathbf{R}^{\mathsf{T}}$).

So that $\mathbf{V}$ corresponds to $\mathbf{S}$, the matrix representation of a discrete sine transform, one must have $\mathbf{S}_{ij} = \mathbf{V}_{ij}$. Since $\mathbf{V}_{ij} = \sin 2x_j\theta_i$, one must have

$$x_j\theta_i = \frac{\pi ij}{N'}, \text{for } i,j = 1,\ldots,M \tag{4.23}$$

Assuming sampling of $d(\theta)$ at sample points

$$\theta_i = \frac{i\theta_u}{M}, \text{ for } i = 1,\ldots,M,$$

this requires

$$x_j = j\Delta x_\circ, \tag{4.24}$$

where $j = 1,\ldots,M$, and $\Delta x_\circ$ is defined:

$$\Delta x_\circ = \frac{\pi M}{N'\theta_u} \tag{4.25}$$

With $x_j = j\Delta x_\circ$, the rows and columns of the matrices $\mathbf{S}$ and $\mathbf{V}$ represent discrete samples of the continuous function $\sin 2x\theta$. Figures 4.1 and 4.2 demonstrate examples of this sampling. Figure 4.3 shows the first several rows of a matrix $\mathbf{G}$, constructed via the diffraction approximation,

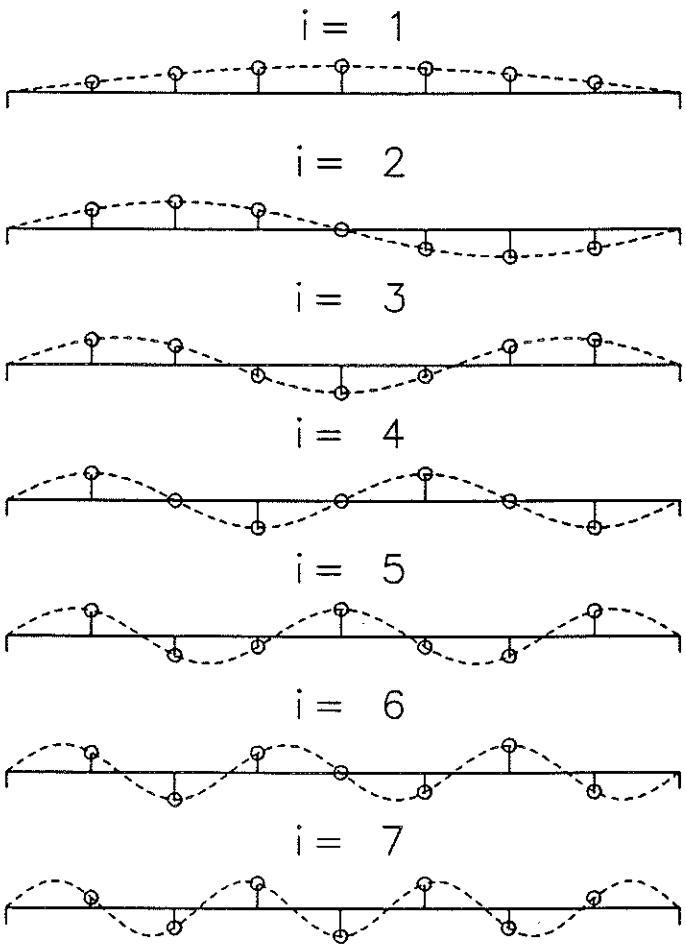

Figure 4.1: $\sin 2x\theta$ (dashed curves) compared with $\sin 2\pi ij/N'$, for $M = 7$

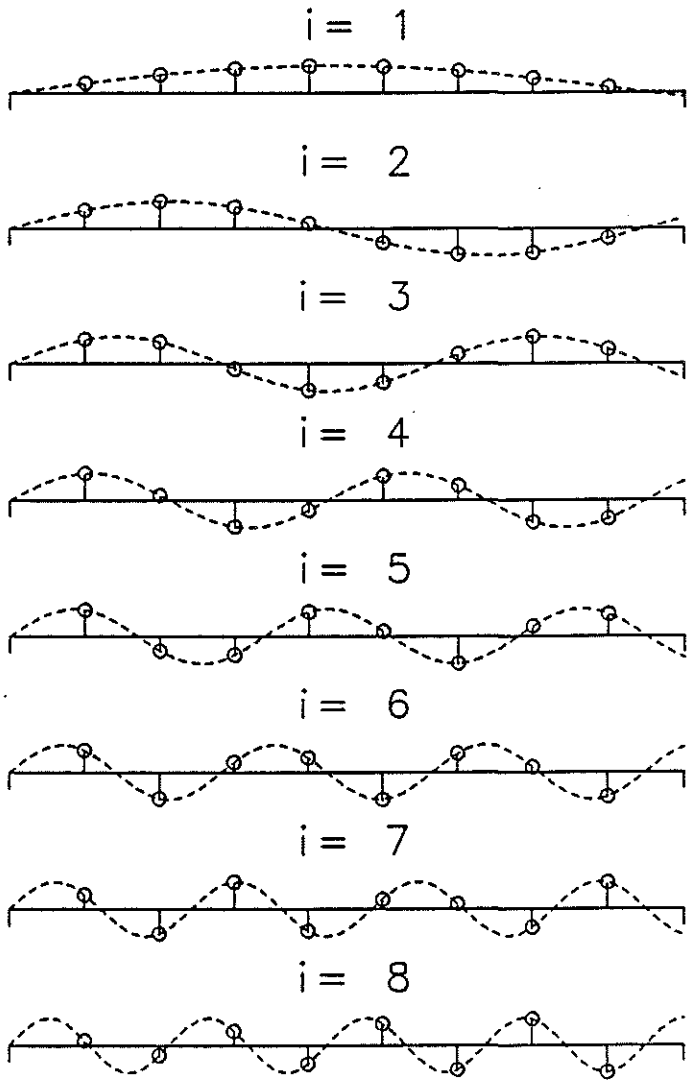

Figure 4.2: $\sin 2x\theta$ (dashed curves) with sample points $\sin 2\pi ij/N'$, for $M = 8$

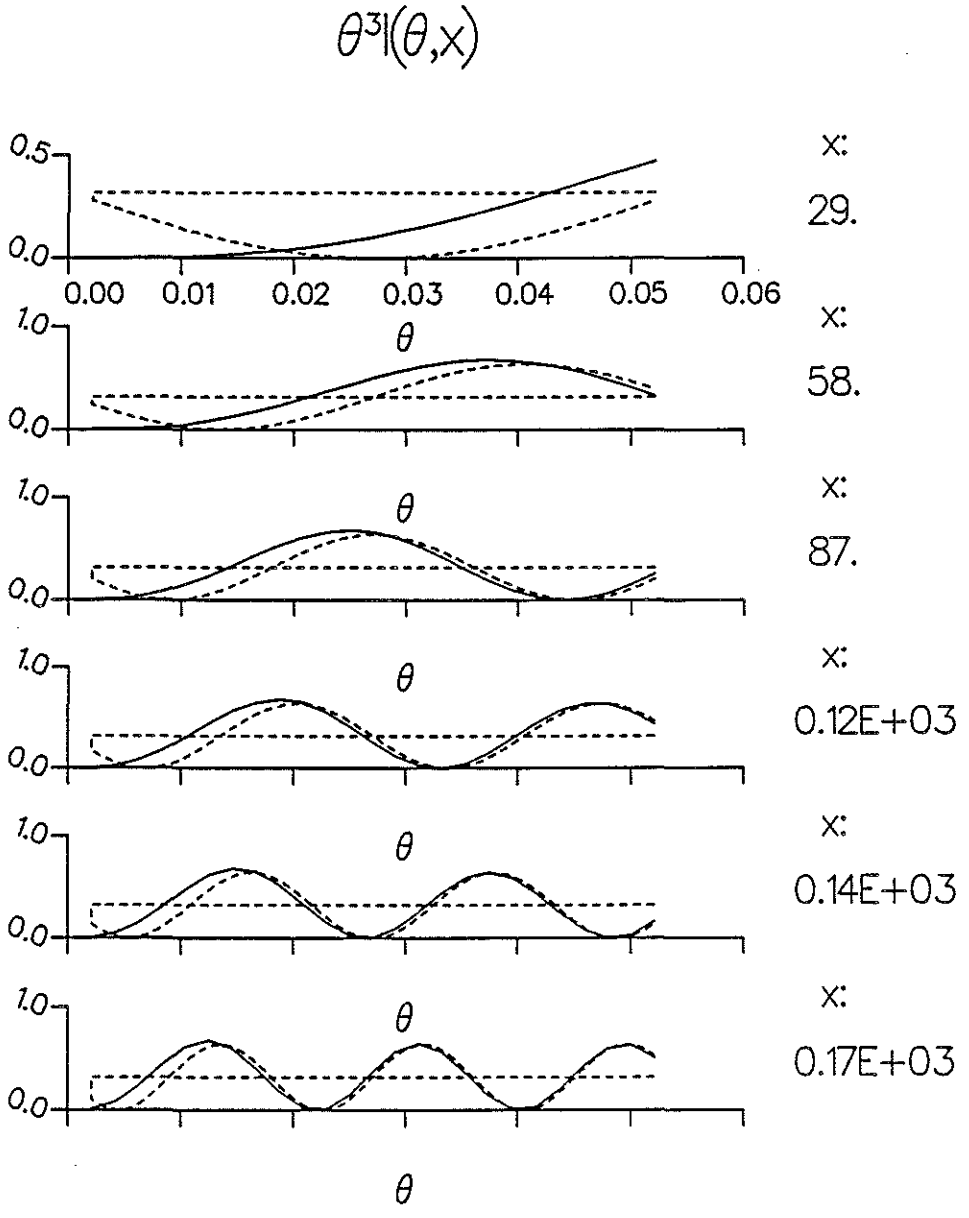

Figure 4.3: The first several rows of a matrix G, constructed via the diffraction approximation (solid lines), plotted with $\frac{1}{\pi k^3}(1-\sin 2x\theta)$ (dashed lines). The matrix was computed using the diffraction approximation. I chose the angular range of θ to be 0.05; the matrix has dimension 25x25.

plotted with the continuous functions $\frac{1}{\pi k^3}(1-\sin 2x\theta)$, and $g_i(x)$ computed from the diffraction approximation.

Having found x_j so that $\mathbf{V}$ has an inverse, the next task is to show that $\mathbf{G}' = \mathbf{U} - \mathbf{V}$ is invertible. The matrix $\mathbf{U}$ is not invertible; this is clear because every element of $\mathbf{U}$ is equal to one. However, since $\mathbf{V}$ is invertible, it is unlikely that adding or subtracting a matrix such as $\mathbf{V}$ will produce an non-invertible matrix, unless $\mathbf{V}$ is particularly pathological. Empirically, I have found that in fact $\mathbf{G}'$ is invertible, though this is usually not the case for $\mathbf{G}$, which $\mathbf{G}'$ approximates.

The task remains to show that when $\mathbf{G}$ is constructed with $x_j = j\Delta x_\circ$, the discrete sums $\mathbf{d} = \mathbf{Gm}$ accurately approximate the inner products $d_i = \langle g_i, m\rangle$. It is evident from fig. 4.3 that $\Delta x_\circ$ is a coarse integration increment. I will examine each term of the approximation

$$\mathbf{d} \approx \frac{\Delta x}{\pi k^3}(\mathbf{Um} - \mathbf{Vm}) \tag{4.26}$$

to which the diffraction approximation is asymptotically equivalent.

The Constant Term The term $\mathbf{Um}$ in (4.26) represents a summation which is approximately the integral

$$[\mathbf{Um}]_i \approx \int_0^{x_{Nyquist}} m(x)\,dx \tag{4.27}$$

The integration increment is small enough if $m(x)$ is reasonably smooth. If $m(x)$ has sharp peaks or discontinuites, this term may be a source of error in inversions. This means that, in addition to the Nyquist criterion, there

is the further restriction that

$$\sum_i m(i\Delta x)\Delta x \approx \int_0^{x_{Nyquist}} m(x)\,dx \tag{4.28}$$

must be a good approximation.

The Periodic Term The term Vm in (4.26) is a summation which is approximately:

$$[\mathrm{Vm}]_i \approx \int_0^{x_{Nyquist}} m(x)\,\sin 2x\theta_i\,dx \tag{4.29}$$

Whether or not Δx_o is small enough requires more care. Again, I rely on basic signal processing concepts. [104,105]. I make use of a theorem concerning a square integrable function $f(\theta)$ and its Fourier transform F(x):

$$f(\theta) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} F(x)\,e^{i\theta x}\,dx \tag{4.30}$$

$$F(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(\theta)\,e^{-i\theta x}\,dx \tag{4.31}$$

Statement 8 *If F_r represents the sampling of $F(x)$*

$$F_r = F(r\Delta x)$$

where r is an integer, and

$$\Delta x \leq \frac{\pi}{2\theta_u},$$

then

$$f(\theta) = \Delta x \sum_{r=-\infty}^{\infty} F(r\Delta x)e^{ir\Delta x\theta} \text{ for } -\frac{\pi}{\Delta x} \leq \theta \leq \frac{\pi}{\Delta x}$$

The dual theorem is also of interest:

Statement 9 *If f_s represents the sampling of $f(\theta)$*

$$f_s = f(s\Delta\theta)$$

where s is an integer, and the sampling of $f(\theta)$

$$f_s = f(s\Delta\theta)$$

where s is an integer, and

$$\Delta\theta \leq \frac{\pi}{2x_u},$$

then

$$F(x) = \Delta\theta \sum_{s=-\infty}^{\infty} f(s\Delta\theta)e^{-irx\Delta\theta} \text{ for } -\frac{\pi}{\Delta\theta} \leq x \leq \frac{\pi}{\Delta\theta}$$

I assume that $f(\theta)$ has compact support in a convenient interval:

$$f(\theta) = 0 \text{ for } \mid \theta \mid \leq M\Delta\theta$$

so that, from theorem 8

$$f(\theta) = \Delta x \sum_{r=-M}^{M} F(r\Delta x)e^{-ir\Delta x\theta} \tag{4.32}$$

If $F(\theta)$ is odd, like $d(\theta)$, this reduces to

$$f(\theta) = \Delta x \sum_{r=1}^{M} F(r\Delta x) \sin r\Delta x\theta \tag{4.33}$$

This is the same summation that **Vm** performs in (4.26). Thus, **Vm** renders an exact approximation to the integral it approximates, if

- $d(\theta)$ has compact support, according to definition 2 of section 3.2.4
- $x_j = j\Delta x_o$, $j = 1, \ldots, m$.

Thus, if $d(\theta)$ is sampled sufficiently, $\mathbf{G}$ constructed with integration increment $\Delta x_\circ$ renders an approximation to $d_i = \langle g_i, m \rangle$ which is accurate as the summation performed by the matrix $\mathbf{U}$.

I emphasize that these results were derived from the diffraction approximation, and may thus have validity at small angles, and nowhere else. Readers interested in choosing the x_j appropriately at larger angles are referred to ref. [94]

The choice of $x_j = j\Delta x_\circ$ may prove too sparse for distributions $m(x)$ which contain abrupt discontinuities or singularities or are relatively largest at small x, at which the sine transform approximation is not valid. This suggests the construction of $\mathbf{G}$ with x_j's that are unevenly spaced. i.e., quite close for small x and at $\Delta x_\circ$ for large x. This results in a larger, rectangular matrix $\mathbf{G}$, which will not be invertible though a *pseudo-inverse* can be found. I have not experimented with unevenly spaced x_j, but suggest that the sampling scheme be based on the expansion for the Bessel function J_1 in equation (9.4.6) of Abramowitz and Stegun [106].

Unevenly Spaced Data θ_i

Generally it is not possible to sample $d(\theta)$ evenly between 0 and θ_u. Dynamic range limitations make measurements near $\theta = 0$ difficult or impossible, and the measurement array may be too short to sample up to θ_u. In this section I assume that the sample points are distributed evenly from θ_{min} to θ_{max}, with $\theta_{min} > 0$ and $\theta_{max} \leq \theta_u$.

I showed in section 3.2.5 that, in principle, $d(\theta)$ can be reconstructed

from M of its samples, if the average sample interval, $\overline{\Delta\theta}$:

$$\overline{\Delta\theta} = \frac{\theta_u}{M-1} \tag{4.34}$$

samples according to the criterion:

$$\overline{\Delta\theta} \leq \frac{\pi}{2x_u} \tag{4.35}$$

Furthermore, $d(\theta)$ can be inverted to form $m(x)$. The basis for this assertion was that $d(\theta)$ and $m(x)$ are *uniquely* represented by the samples (except for $m(x)$ at $x = 0$.) Of course, the inversion is unstable; this was discussed in section 3.2.6.

To construct G for inverting unevenly spaced data, I suggest that the points x_j be chosen as before, in terms of the angle θ_u; that is choose $x_j = j\Delta x_\circ$, where $j = 1, \ldots, M$, and

$$\Delta x_\circ = \frac{\pi M}{N'\theta_u}. \tag{4.36}$$

This matrix is the same as constructed for evenly spaced data, except that the sample points θ_i have been migrated to new points θ_i'.

Initially, I suggest choosing $\theta_u = \theta_{max}$. When $d(\theta)$ does not approach an asymptote with in the range of measurements, this choice will not be appropriate. It may, however, suggest bounds on x, in the form of a minimum and maximum size, x_{min} and x_{max}, with $x_{max} < x_{Nyquist}$. Then, a matrix can be chosen with the sample points x_j spaced evenly from x_{min} to x_{max}. Such a matrix corresponds to choosing a larger θ_u, which falls beyond the range of measurements.

4.2 Inversion Algorithms

We have seen how to construct the matrix G so as to reduce its singularity, and the time to compute it, while ensuring that it is physically complete. Next, I consider algorithms for inverting G to find various estimates $\hat{m}(x)$. For our purposes, inversion algorithms can be characterized according to

- treatment of eigenvectors with small eigenvalues,
- the tradeoff between resolution and variance, and
- the constraints on the solution $\hat{m}(x)$, and how these are imposed algorithmically.

Ordinarily, the treatment of eigenvectors with small eigenvalues would not warrant such special attention. For diffraction particle sizing, however, the eigenvectors have significance different from those of many other inverse problems. I discuss this treatment first.

4.2.1 Treatment of Eigenvectors with Small Eigenvalues

The treatment of eigenvectors with small eigenvalues is important in the diffraction particle sizing problem, because it affects the performance of some common algorithms for "inversion" of singular matrices. Some algorithms which work very well in other inverse problems fail for the diffraction particle sizing problem, particularly for distributions with skewed toward the small sizes.

A square MxM matrix $\mathbf{Q}$ can be decomposed in the form

$$\mathbf{Q} = \mathbf{A}\Lambda\mathbf{A}^{\mathsf{T}} \tag{4.37}$$

where Λ is a pxp matrix with the p non-zero eigenvalues, $\lambda_1, \ldots, \lambda_p$ of $\mathbf{Q}$ along its diagonal, ordered so that $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_N$. The MxM matrix $\mathbf{V}$ contains the eigenvectors of $\mathbf{Q}$ as columns. For MxN rectangular matrices this may be generalized: [85]

$$\mathbf{Q} = \mathbf{A}\Lambda\mathbf{B}^{\mathsf{T}} \tag{4.38}$$

where Λ contains the eigenvalues determined by the generalized eigenvector equations

$$\mathbf{Q}\mathbf{a}_i = \lambda_i \mathbf{b}_i \tag{4.39}$$

$$\mathbf{Q}^{\mathsf{T}}\mathbf{b}_j = \lambda_j \mathbf{a}_j \tag{4.40}$$

where $\mathbf{a}_i$ and $\mathbf{b}_j$ are the eigenvectors that make up the columns of $\mathbf{A}$ and $\mathbf{B}$ respectively. The matrix $\mathbf{Q}$ is considered ill-conditioned for inversion when the ratio λ_1/λ_N is large. Singularity is associated with the smallest eigenvalues. The eigenvalues serve as weighting factors of the eigenvectors, which can be regarded as bearers of information. The matrix $\mathbf{Q}$ tends to attenuate information contained in the eigenvectors associated with small eigenvalues. In the optical particle sizing problem, there is little information in $d(\theta)$ regarding $m(x)$ at small x. Hence, the eigenvectors of $\mathbf{G}$ corresponding to small eigenvectors tend to represent the small particles in $d(\theta)$. This can be seen in 4.4, a plot of the rows of the matrix $\mathbf{B}$. The eigenvector associated with the smallest eigenvalue is largest at the smallest

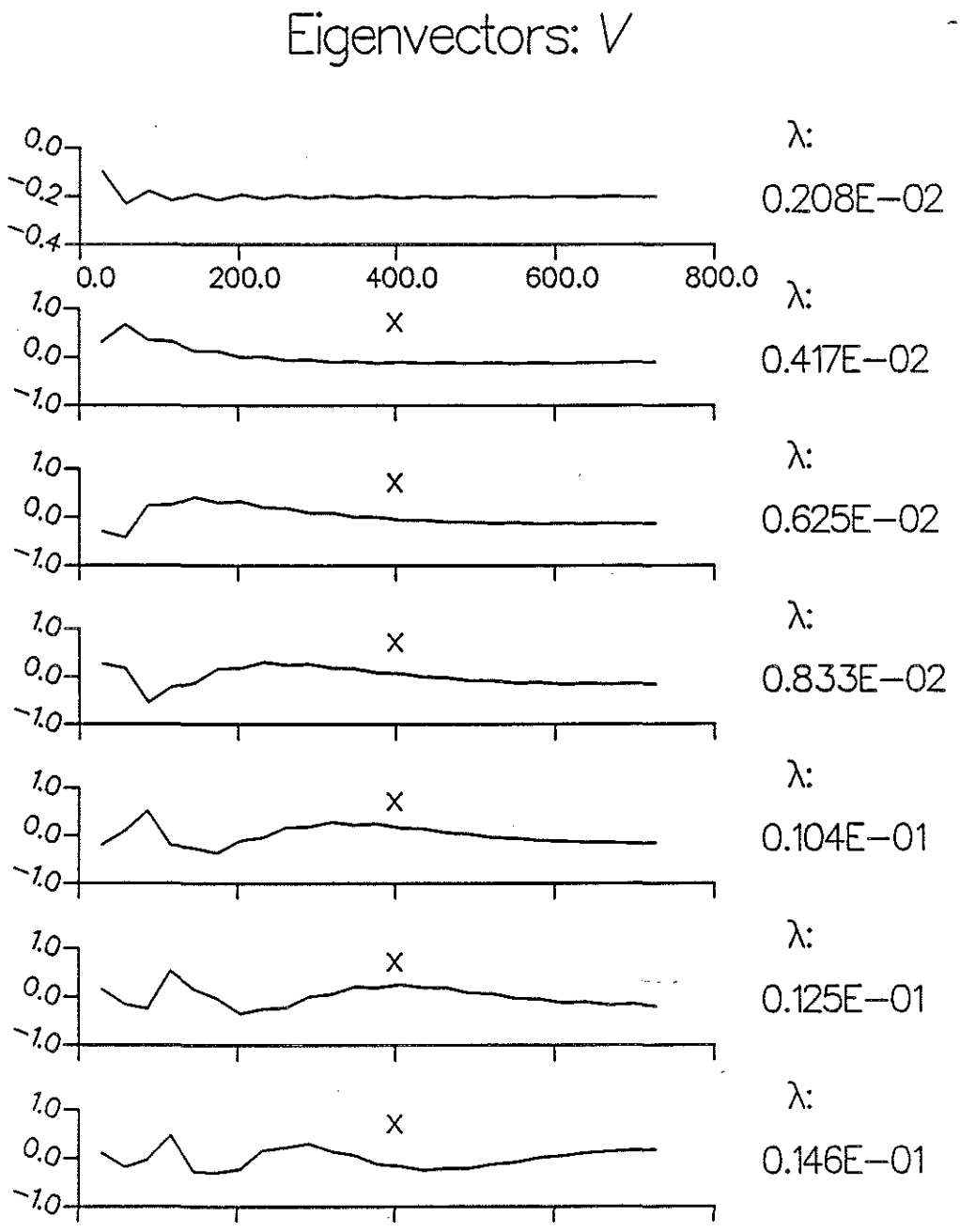

Figure 4.4: A plot of some of the eigenvectors of G. This is the same matrix as in figure 4.3.

size x. Most of the larger eigenvalues are nearly equal, with the result that the corresponding eigenvectors are not unique; hence these have random shape.

The eigenvectors in the optical particle sizing problem serve a different role from those in other inverse problems. In most problems, $\mathbf{G}$ is at least crudely a low pass filter. Small eigenvalues are associated with high frequencies in the model m, which $\mathbf{G}$ attenuates most. The associated eigenvectors have the most high frequency content. By contrast, $\mathbf{G}$ in the optical particle sizing problem is a high pass filter, which passes all frequencies equally well except for the very lowest.

The role of the eigenvectors and eigenvalues has important implications for *pseudo-inverses* $\mathbf{H}$ to $\mathbf{G}$; many pseudo-inverses can be characterized by the treatment of the eigenvalues. One, for example, is the "Lanczos" inverse. [85,107,108]. In this algorithm the eigenvectors associated with small eigenvalues are simply eliminated from the *singular value decomposition*:

$$\mathbf{G} = A\Lambda B^{\mathsf{T}} \tag{4.41}$$

where $\mathbf{G}$ has dimensions MxN. Thus an estimate $\hat{\mathbf{m}}$ is formed from the sum

$$\hat{\mathbf{m}}_j = \sum_{k=1}^{q} \sum_{i=1}^{M} \mathbf{B}_{ik} \lambda_k^{-1} \mathbf{A}_{jk} \mathbf{d}_i, \tag{4.42}$$

where the summation over eigenvectors is terminated at an index $q < p$ rather than at p.

Another common algorithm for $\mathbf{H}$ which can be characterized by its effect on the eigenvalues as the Damped Least Squares inverse.[109,110] In

the inverse $\mathbf{H}$ is written as

$$\mathbf{H} = \begin{array}{ll} \mathbf{G}^{\mathsf{T}}(\mathbf{G}\mathbf{G}^{\mathsf{T}} + \epsilon\mathbf{I})^{-1} & \text{for } M \leq p \leq N \\ (\mathbf{G}^{\mathsf{T}}\mathbf{G} + \epsilon\mathbf{I})^{-1}\mathbf{G}^{\mathsf{T}} & \text{for } M \geq p \geq N \end{array} \tag{4.43}$$

In either case $\mathbf{H}$ can be decomposed via singular value decomposition in the form $\mathbf{H} = \mathbf{B}\mathbf{\Lambda}^{-1}\mathbf{A}^{\mathsf{T}}$ where

$$\Lambda_{ij}^{-1} = \frac{\delta_{ij}\lambda_i}{\lambda_i^2 + \epsilon} \tag{4.44}$$

δ_{ij} is the Kronecker delta and the λ_i are the eigenvalues of $\mathbf{G}$. Here, the eigenvectors with small eigenvectors are suppressed by modifying the eigenvalues with the constant ϵ.

In most inverse problems where small eigenvalues correspond to eigenvectors with high frequency content, the effect of these inverses is to smooth the model $\hat{\mathbf{m}}$ by filtering out high frequencies. In the optical particle sizing problem, however, the effect is to corrupt $\hat{\mathbf{m}}$ at small sizes. This has serious consequences for size distributions prevalent in nature. In many settings, as we saw in section 1.3, $m(x)$ is best represented by a power law of the form

$$m(x) = Cx^{-\tau} \text{ for } x > x_0 \tag{4.45}$$

where $\tau \approx 3$ or 4. Manipulating small eigenvalues could severely distort this size distribution.

The essential result of this section is that algorithms which are *not* based on manipulating small eigenvalues should be chosen for inverting $d(\theta)$ when distortion at small sizes is a problem.

4.2.2 The Resolution-Variance Tradeoff

The results of any inversion algorithm can be characterized by resolution, the ability to resolve peaks, and variance, the sensitivity to noise in the data. Some inversion algorithms actually allow an investigator to *choose* the tradeoff. Two such algorithms are the Backus-Gilbert and Phillips-Twomey algorithms.

The Formalism of Backus and Gilbert The Backus–Gilbert formalism is ideal for examining the tradeoff between resolution and variance, because, unlike other algorithms, it treats resolution and variance explicitly. Unfortunately, the formalism is inefficient computationally. Hence, I present the formalism for examining the tradeoff, and not for actually inverting for $m(x)$. In principle the formalism yields inversions that have an optimal resolution–variance product. The formalism produces intractable algebra for the optical particle sizing problem, however. Fortunately, it can yield results for the similar problem of determining Fourier transforms. I review the Backus–Gilbert formalism in the following paragraphs. The results support section 3.64, in which I stated the best resolution obtainable for $\hat{m}(x)$.

I have written the data d_i in terms of the model $m(x)$ as

$$d_i = \int_0^\infty m(x) g(\theta_i, x)\, dx \tag{4.46}$$

For a specific size x_0 an estimate $\hat{m}(x_0)$ can be written in terms of a linear combination of the data:

$$\hat{m}(x_0) = \sum_{i=1}^{M} \alpha_i(x_0) d_i \tag{4.47}$$

where the coefficients $\alpha_i(x_0)$ are functions of the specific, "target" size x_0. Substituting (4.46) into (4.47) yields

$$\hat{m}(x_0) = \int_0^\infty m(x) \sum_{i=1}^{M} \alpha_i(x_0) g(\theta_i, x)\, dx \tag{4.48}$$

at which point it is convenient to define

$$\hat{\delta}(x, x_0) = \sum_{i=1}^{M} \alpha_i(x_0) g(\theta_i, x). \tag{4.49}$$

Ideally one would have $\hat{\delta}(x, x_0) = \delta(x - x_0)$, where $\delta(x - x_0)$ is the Dirac delta function, so that equation (4.48) would become $\hat{m}(x_0) = m(x_0)$. Of course, this can happen only if M is infinite.

To find $\hat{m}(x_0)$ we need a set of coefficients $\alpha_i(x_0)$. We will require the normalization

$$\int_0^\infty \hat{\delta}(x, x_0)\, dx = 1 \tag{4.50}$$

or

$$\sum_{i=1}^{\infty} \alpha_i(x_0) \int_0^\infty g(\theta_i, x)\, dx = 1, \tag{4.51}$$

which can be written in the convenient matrix form

$$\alpha^{\mathsf{T}} \mathrm{u} = 1. \tag{4.52}$$

Here $\alpha^{\mathsf{T}} = [\alpha_1(x_0), \alpha_2(x_0), \ldots, \alpha_M(x_0)]$ and

$$u_i = \int_0^\infty g(\theta_i, x)\, dx. \tag{4.53}$$

The $\alpha_i(x_0)$ are found by minimizing a measure of the "closeness" of $\hat{\delta}(x, x_0)$ to the Dirac delta function. Three measures are common in the

literature[1]:

$$D_1(x_0) = \int_0^\infty [\hat{\delta}(x, x_0) - \delta(x - x_0)]^2 \, dx \tag{4.54}$$

$$D_2(x_0) = 12 \int_0^\infty \left[\int^x \hat{\delta}(x', x_0) \, dx' - H(x - x_0) \right]^2 dx \tag{4.55}$$

and

$$S(x_0) = 12 \int_0^\infty (x - x_0)^2 \hat{\delta}^2(x - x_0) \, dx \tag{4.56}$$

where $H(x - x_0)$ is the Heaviside step function.[2] The functions $D_1(x_0)$ and $D_2(x_0)$ are known as the first and second Dirichlet criteria, respectively; $S(x_0)$ is known as the *spread.* The factors of 12 are for normalization.

Of course, noise sensitivity (or variance) is as important as resolution. Suppose that the array $\mathbf{n}^\mathsf{T} = [n_1, n_2, \ldots, n_M]$ is an array of additive noise which contaminates the data d_i. The noise has a covariance matrix

$$\mathrm{C} = \langle \mathbf{n}\mathbf{n}^\mathsf{T} \rangle \tag{4.57}$$

where the brackets denote the expectation operator. The variance of the estimate $\hat{m}(x_0)$ is

$$\alpha^\mathsf{T} \mathrm{C} \alpha$$

and this must be minimized.

---

[1] The lower limit of integrations in the literature is always $-\infty$; here I specialize to the particle sizing problem where $x > 0$.

[2] The Heaviside step function is defined as

$$H(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1/2 & \text{if } x = 0 \\ 1 & \text{for } x > 0 \end{cases}$$

Unfortunately, it is not possible to minimize (4.54), (4.55), or (4.56) and the variance simultaneously. In general there is a tradeoff. This is most easily seen when the spread (of equation (4.56)) and the variance are minimized together. The spread can be written in matrix form

$$S(x_0) = \alpha^{\mathsf{T}} \mathrm{R} \alpha \tag{4.58}$$

where

$$R_{ij} = \int_0^\infty (x - x_0)^2 g(\theta_i, x) g(\theta_j, x)\, dx, \tag{4.59}$$

which is exactly in the same form as the variance. In practice a factor of merit

$$\begin{aligned} \mathcal{F} &= q\alpha^{\mathsf{T}} \mathrm{R} \alpha + (1-q)\alpha^{\mathsf{T}} \mathrm{C} \alpha \\ &= \alpha^{\mathsf{T}} [q\mathrm{R} + (1-q)\mathrm{C}] \alpha \end{aligned} \tag{4.60}$$

is minimized, subject to the constraint $\alpha^{\mathsf{T}} \mathbf{u} = 1$. Here q is a parameter between 0 and 1. Note that $q = 0$ results in minimum variance, while $q = 1$ results in maximum resolution. The minimization is easy with the method of Lagrange multipliers; the result is

$$\alpha = \left[\mathbf{u}[q\mathrm{S} + (1-q)\mathrm{C}]^{-1}\mathbf{u}^{\mathsf{T}}\right] [q\mathbf{S} + (1-q)\mathbf{C}]^{-1}\mathbf{u} \tag{4.61}$$

To study the tradeoff between resolution and variance for the optical sizing problem, one would ideally substitute the kernel $G(\theta, x)$ into the formalism, and obtain values for the spread or the Dirichlet criteria. Setting $q = 0$, the criteria would indicate the best resolution obtainable. The equations, however are intractable for this inverse problem. Fortunately,

because $m(x)$ and $d(\theta)$ are approximately related by a Fourier-sine transform, the Backus-Gilbert formalism is not necessary – the best obtainable resolution is well known from the theory of Fourier transforms. This was discussed in section 3.3.

Oldenberg [111] examined the resolution–variance tradeoff in the Backus-Gilbert formalism, for Fourier transforms. His work is applicable to the diffraction particle sizing. Not surprisingly, Oldenberg finds that when the first Dirichlet condition of equation (4.54) is minimized, and the tradeoff is chosen to provide maximum resolution, the result is identically the discrete Fourier transform. The best resolution possible is the same as a discrete Fourier transform (DFT) would yield. So, as expected, the DFT yields a maximum resolution rather than minimum variance estimate. For this work, the important conclusion is that the resolution of equation (3.64) is the best resolution possible, and is supported by the Backus-Gilbert formalism.

The Phillips-Twomey Algorithm The Phillips-Twomey algorithm is a least squares algorithm that has an additional constraint — that the estimate $\hat{\mathbf{m}}$ be smooth. Necessarily, resolution suffers from the constraint of smoothness, but the estimate $\hat{\mathbf{m}}$ is less sensitive to noise in the data. Thus, the effect is to trade sensitivity to noise for resolution, much the same as for Backus-Gilbert, except that the tradeoff is not optimal in the same sense.

Clearly there are many measures of the "smoothness" of a function. The principle of the Phillips-Twomey algorithm is to find a model with a

minimum second derivative. The second derivative in discrete form of $\hat{\mathbf{m}}$ can be written in matrix form as $\mathbf{K}\hat{\mathbf{m}}$, where

$$\mathbf{K} = \begin{pmatrix} 0 & 0 & . & . & . & . & . \\ -1 & 2 & -1 & . & . & . & . \\ . & -1 & 2 & -1 & . & . & . \\ . & . & -1 & 2 & -1 & . & . \\ . & . & . & -1 & 2 & -1 & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & 0 \end{pmatrix} \tag{4.62}$$

The matrix $\mathbf{K}$ approximates a second derivative by finding second differences. The measure of smoothness is then $\hat{\mathbf{m}}^T\mathbf{K}^T\mathbf{K}\hat{\mathbf{m}}$, which is the discretized integral of the second derivative of $\hat{\mathbf{m}}$, squared. The term $\mathbf{K}^T\mathbf{K}$ may be rewritten as

$$\mathbf{K}^T\mathbf{K} = \begin{pmatrix} 1 & -2 & . & . & . & . & . & . \\ -2 & 5 & -4 & 1 & . & . & . & . \\ 1 & -4 & 6 & -4 & 1 & . & . & . \\ . & 1 & -4 & 6 & -4 & 1 & . & . \\ . & . & 1 & -4 & 6 & -4 & 1 & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & -2 & 1 \end{pmatrix} \tag{4.63}$$

The Phillips-Twomey algorithm minimizes $\hat{\mathbf{m}}^T\mathbf{K}^T\mathbf{K}\hat{\mathbf{m}}$ subject to the least squares constraint

$$(\mathbf{G}\hat{\mathbf{m}} - \mathbf{d})^T(\mathbf{G}\hat{\mathbf{m}} - \mathbf{d}) \leq \epsilon \tag{4.64}$$

where ϵ is the least square error, or "missfit." This is a straight forward minimization, which can be performed by the method of Lagrange multipliers. In this method, the derivatives of

$$(\mathbf{G}\hat{\mathbf{m}} - \mathbf{d})^{\mathsf{T}}(\mathbf{G}\hat{\mathbf{m}} - \mathbf{d}) + \nu\hat{\mathbf{m}}^{\mathsf{T}}\mathbf{K}^{\mathsf{T}}\mathbf{K}\hat{\mathbf{m}}$$

with respect to all elements $\hat{m}_j$ and the Lagrange multiplier ν are set to zero. The result is

$$\hat{\mathbf{m}} = (\mathbf{G}^{\mathsf{T}}\mathbf{G} + \nu\mathbf{K}^{\mathsf{T}}\mathbf{K})^{-1}\mathbf{G}^{\mathsf{T}}\mathbf{d} \tag{4.65}$$

The effect of the added constraint of smoothness is to reduce resolution, but decrease sensitivity to noise. Of course, there may be size distributions $m(x)$ for which the smoothness constraint is inappropriate, i.e. distributions which have peaks or are otherwise steep. An example is the Junge distribution of section 1.3.

4.2.3 Algorithms Which Impose Other Constraints

All of the algorithms discussed thus far are, in some sense, least squares algorithms. That is, all produce an estimate $\hat{m}(x)$, subject to the constraint that either a least square error is minimized, or the estimate has least square-norm. Other algorithms impose different, or additional constraints. A number of iterative algorithms exist, which can constrain $\hat{m}(x)$ to be positive. Linear programming techniques, for example, minimize least *absolute* magnitude.

Iterative Algorithms

Iterative algorithms are especially useful for imposing *a priori* knowledge in the form of non-linear constraints in inversions. For example, a solution can be constrained to be band-limited, as in extrapolation and super-resolution problems. We are particularly interested in constraining the estimate $\hat{m}(x)$ to be positive; this is most useful when scattering measurements are not available at small angles.

A convenient way to express *a priori* knowledge in the form of a constraint, is as an operator, C, so that

$$\mathbf{m} = \mathbf{Cm} \tag{4.66}$$

if $\mathbf{m}$ satisfies the constraint. The operator C can impose band-limitedness, if C is a filter, or positivity, as in the optical particle sizing problem. The forward problem, $\mathbf{d} = \mathbf{Gm}$ can then be written

$$\mathbf{d} = \mathbf{GCm} \tag{4.67}$$

One class of iterative algorithms is based on forming the trivial identity

$$\mathbf{m} = \mathbf{Cm} + \lambda(\mathbf{d} - \mathbf{GCm}) \tag{4.68}$$

from equations (4.66) and (4.67).[112] Here λ can be any function, matrix, or perhaps a constant. One example is Landweber iteration, which can be written in the form

$$\hat{\mathbf{m}}^{k+1} = \mathbf{C}\hat{\mathbf{m}}^{k} + \alpha\mathbf{G}^{\mathsf{T}}(\mathbf{d} - \mathbf{GC}\hat{\mathbf{m}}^{k}) \tag{4.69}$$

where $\hat{\mathbf{m}}^{k}$ is the estimate of $\mathbf{m}$ at the k-th iteration, and α is an arbitrary but small (< 1) damping constant for increasing stability [113]. With no

constraint C, it can be shown that this iteration converges to the least squares result, $\hat{\mathbf{m}} = (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\mathbf{d}$. In the limit of infinite iterations, then, this algorithm is no more stable than any least squares algorithm. It can have good stability however, after a finite number of iterations. Later, will compare this algorithm with others when I present inversions of experimental data.

Schafer reviews iterative algorithms, and disccusses the convergence of iterative algorithms based on equation (4.68) in reference [112]. The proof of convergence can be difficult, and is beyond the scope of this work. It can sometimes be proven by writing the iteration in the form

$$\hat{\mathbf{m}}^{k+1} = \mathbf{F}\hat{\mathbf{m}}^k \tag{4.70}$$

where $\mathbf{F}$ is an operator, and showing that $\mathbf{F}$ is a *contraction mapping.* This means that for arbitrary models $\mathbf{m}_i$ and $\mathbf{m}_j$, the inequality

$$||\mathbf{F}\mathbf{m}_i - \mathbf{F}\mathbf{m}_j|| \leq ||\mathbf{m}_i - \mathbf{m}_j|| \tag{4.71}$$

applies, where the norms are the Euclidean norm. In practice, it may be possible to show that an operator $\mathbf{F}$ is a contraction operator in the continuous case, but not in the discrete case. In my experience, the numerical round-off, or the setting extremely small numbers to zero in a computer, adds a further complication, so that in practice $\mathbf{F}$ may not be a contraction mapping after all.

A second iterative algorithm based on (4.68) I call the *projection iteration* inversion. [114] In this method, each iteration is part of a larger cycle.

In each iteration,

$$\hat{\mathbf{m}}^{k+1} = \mathbf{C}\left[\hat{\mathbf{m}}^k - \frac{[\hat{\mathbf{m}}^k \cdot \mathbf{g}_i - \mathbf{d}_i]}{\mathbf{g}_i \cdot \mathbf{g}_i}\mathbf{g}_i\right], \tag{4.72}$$

is cycled over i, which corresponds to the measurement angles. Here $\mathbf{g}_i$ is a vector made from the i-th row of $\mathbf{G}$, and $\cdot$ represents the vector dot product. Like Landweber iteration, this algorithm also converges to the least squares solution. Unfortunately this algorithm is computationally intensive.

A third iterative algorithm is not based on equation (4.68). I call this the modified Chahine algorithm, since it is based on an algorithm by Chahine[113,88]. The iterations of the algorithm are defined by

$$\hat{\mathbf{m}}_j^{k+1} = \mathbf{C}[1 + \xi_1 G_{1j}][1 + \xi_2 G_{2j}] \ldots [1 + \xi_M G_{Mj}]\hat{\mathbf{m}}_j^k \tag{4.73}$$

where

$$\xi_i = \frac{d_i}{(\mathbf{G}\hat{\mathbf{m}}^k)_i} \tag{4.74}$$

In practice it is desirable to restrict the ξ_i to fall with in an interval, $\upsilon \leq \xi_i \leq \upsilon$. As the iteration proceeds this interval should be reduced. The rate of convergence depends on the history of changes to υ during the iteration. A potential defect of this algorithm is that once any $\hat{m}_j$ equals zero, it can no longer change, since all iterates are multiplied by previous iterates. In a computer where all numbers smaller than some minimum are set to zero, this can eventually cause convergence to fail, or convergence to an erroneous result.

A fourth iterative algorithm is the non-negative least squares (NNLS) algorithm described by Lawson and Hanson [115], or the positive least squares (PLS) algorithm described by Menke [116]. These are essentially the same

algorithm, which purport to find the solution to an over-determined system $\mathbf{Gm} = \mathbf{d}$ with the constraint that all the unknowns are non-negative. A description of this algorithm is beyond the scope of this work, particularly since it gives disappointing results with experimental data.

Linear Programming

I have not inverted data using linear programming, but I discuss it for completeness, and reference for future work. The standard linear program [117,118] is to minimize a penalty function

$$z = \mathbf{c}^T\mathbf{m} \tag{4.75}$$

where $\mathbf{c}$ is a "cost" vector[3], subject to the constraints

$$m_j = 0 \tag{4.76}$$

with $j = 1, \ldots, N$ and

$$\mathbf{d} = \mathbf{Gm}. \tag{4.77}$$

The solution of a linear program is performed using any of several readily available algorithms, such as the Simplex algorithm.

There is a further restriction that (4.77) represent an under-determined set of equations, that is, that $N \gg M$, recalling that $\mathbf{G}$ is an NxM matrix. This restriction makes linear programming less interesting for our work, where M typically represents about 100 points of experimental data. This

[3]The cost vector is often chosen to have all elements equal to 1, as in a minimization problem, or -1, as in a maximization problem.

means that must be at least 100x100, requiring considerable computational time.

The standard problem is easily transformed to solve several related problems, so that linear programming has some appeal for future work. For example, instead of equation (4.77) one might desire the constraint

$$\mathbf{Gm} \leq \mathbf{d}_u \tag{4.78}$$

where $\mathbf{d}_u$ is a vector representing an upper bound. This constraint can be accomdated through the addition of a vector of slack variables, $\mathbf{y}$, so that

$$\mathbf{G}\hat{\mathbf{m}} + \mathbf{I}\mathbf{y} = \mathbf{d}_u. \tag{4.79}$$

where is the identity matrix. For (4.77) to apply, we must have $\mathbf{Gm} \leq \mathbf{d}_u$. It is convenient to define two new vectors

$$\mathbf{w} = \begin{pmatrix} \mathbf{m} \\ \mathbf{y} \end{pmatrix} \tag{4.80}$$

and

$$\mathbf{c}' = \begin{pmatrix} \mathbf{c} \\ \mathbf{0} \end{pmatrix} \tag{4.81}$$

where $\mathbf{0}$ represents a vector of zeros, and a matrix

$$\mathbf{B} = \begin{pmatrix} \mathbf{G} & \mathbf{I} \end{pmatrix} \tag{4.82}$$

With these definitions, one can rewrite the linear program as a standard linear program. Thus, the penatly function is

$$z = \mathbf{c}'^{\mathsf{T}}\mathbf{z} \tag{4.83}$$

subject to the constraints

$$z_i \geq 0 \tag{4.84}$$

and

$$\mathbf{Bw} = \mathbf{d} \tag{4.85}$$

Similarly it is possible to write a linear program to satisfy the constraint

$$\mathbf{d}_l \leq \mathbf{Gm} \leq \mathbf{d}_u \tag{4.86}$$

where $\mathbf{d}_l$ and $\mathbf{d}_u$ represent lower and upper bounds on $\mathbf{d}$, respectively. This constraint may prove useful in inverting noisey data, particularly where digitization noise is severe.

4.3 Summary

Starting with an asymptotic analysis of the diffraction approximation, I have reached the following conclusions regarding inversions of near forward scattered data:

- A Nyquist criterion applies to the measurements of $I(\theta)$, for a given size distribution $n(x)$.
- For many $n(x)$, $\theta^3 I(\theta)$ approaches a constant. This justifies, in part, sampling only at near forward angles. Moreover, it allows for a more generous Nyquist criterion.
- It is numerically much more stable to invert $d(\theta) = \theta^3 I(\theta)$ to yield $m(x) = xn(x)$ rather than $I(\theta)$ to yield $n(x)$, because the kernel

$G(\theta, x) = \theta^3 I(\theta, x)/x$ is of constant order, while $I(\theta, x)$ varies over orders of magnitude.

- For robust inversion, $d(\theta)$ (and equivalently $I(\theta)$) should be sampled evenly from $\theta = 0$ to θ_u. Sampling from θ_{min} to θ_{max} results in instability. This is not severe if θ_{min} is small and θ_{max} is close to θ_u. If this instability can be surmounted, it would provide a way to measure the volume scattering function, by extrapolating measurements at larger angles.

- The asymptotic Fourier transform relation between $d(\theta)$ and $m(x)$ can be exploited to increase the efficiency of computing one from the other, by showing how to choose the coarsest grid for integration. This grid is fine enough if the Nyquist criterion is satisfied, and if $\int m(x)dx \approx \sum m(i\Delta x)\Delta x$ is a good approximation. A coarse grid for integration allows inversions of data with over-determined matrices.

- Matrix inversion algorithms show promise, but algorithms which explicitly or implicitly manipulate small eigenvalues may corrupt distributions at small sizes; this is a result of mathematical instability at small sizes.

Chapter 5

Experimental Measurements of Small Angle Light Scattering

The theory of the previous chapters has driven my experimental program. While my goal is to develop an instrument for sizing particle size distributions in oceanic boundary layer flows, the theory has general applicability to, for example, the sizing of aerosol sprays.

In this work, I use experimental data to test the theory of the previous chapters. There are several reasons for using experimental data rather than, or in addition to, synthetic data. First, the experimental data is contaminated by sources of noise which are not understood well enough to model. For the study of many inverse problems, investigators model noise as random, white and Gaussian. Here, however, such noise is easily reduced by averaging. To model it is essentially pointless, since other, non-random noise, dominates. Second, there is a tendency among investigators to learn how to generate the synthetic data which yield the best results, and then use this in all their work, even if it is not realistic data. Third, difficulties

which do not arise in synthetic data manifest in experimental data. For example, missing data need not occur in synthetic data, but it certainly is a problem experimentally.

The mathematical instability associated with small particles and the instability associated with less than robust sampling causes the inversion of $d(\theta)$ to be sensitive to noise. To some degree this sensitivity can be tempered by adding constraints to inversion algorithms, for example, insisting that the size distribution found be smooth or positive. The best way to reduce instability, however, is to obtain scattering measurements evenly spaced from $\theta = 0$, but not including $\theta = 0$, to θ_u.

Obtaining such measurements of scattering is difficult, however, for two reasons:

- The dynamic range of the scattered beam $I(\theta)$ is large, partly because the kernel $I(\theta, x)$ is of order $x/(k\theta^3)$ at near forward angles.
- The intense, unscattered beam intrudes, which by fault of its Gaussian profile, requires increased dynamic range.

5.1 Small Angle Scattering Measurements: Collecting the Data

Here I describe in detail a prototype instrument for particle sizing, which I have used to gather data for testing the theory of prior chapters, on measurements from distributions of particles. The instrument is based on a CCD photodiode array. The instrument acheives sufficient dynamic

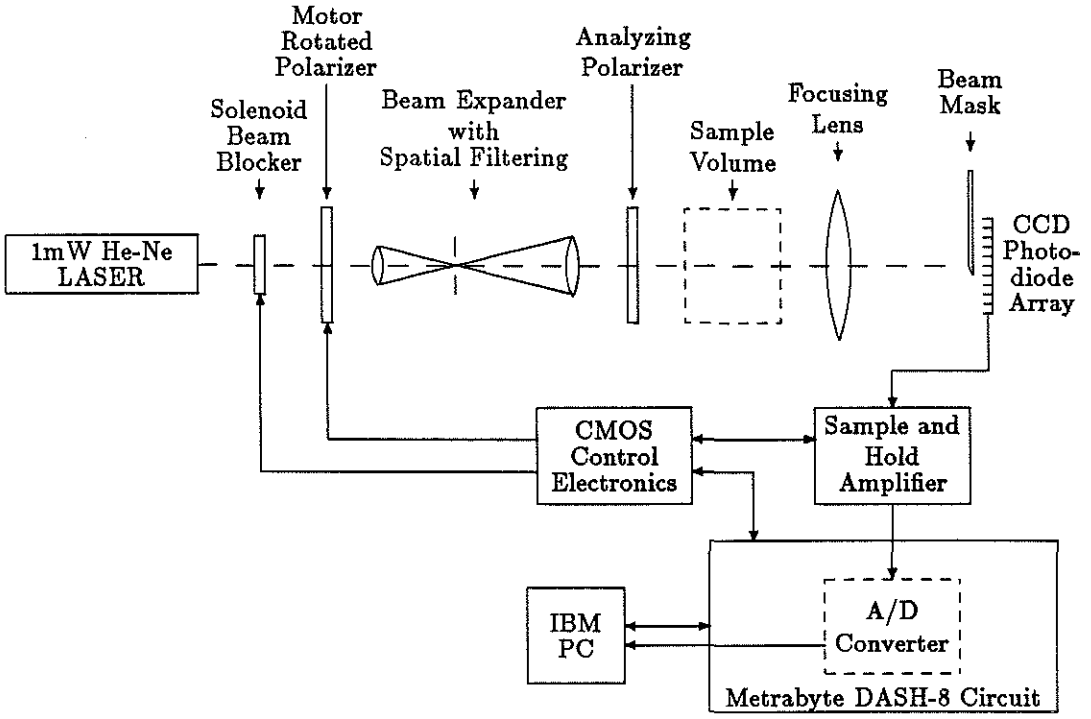

Figure 5.1: Schematic of the optical particle sizing instrument.

range by computer averaging many sets of measurements. It overcomes the intrusion of the unscattered beam by use of a knife edge mask, overlying the photodiode array.

Figure 5.1 is a schematic of the instrument. In the next paragraphs, I discuss each element of the instrument, in the order in which light passes through the instrument, and in the order in which signals pass through the electronics.

Laser The laser is a Hughes model number 3221H-P. This laser has a minimum CW output power of 1 milliwatt at 632.8 nm in the TEM_{00} mode. The beam diameter at the $1/e^2$ power points is 0.64 mm. The output is polarized.

Intensity controller In order to adjust light levels so that measured intensities fall within the dynamic range of the detector, I constructed an intensity controller consisting of a polarizing filter rotated by a stepping motor, which is under computer control. Crossed polarizers can control beam intensity over a range of almost 10^5, allowing for greatly increased dynamic range. The stepping motor is a Hurst Mfg. part number 3004-004, with gear reduction to give a 0.25° step angle. With gearing, the step angle for the polarizer is 0.04545°, so that a rotation of 1000 steps moves the polarizer through one quarter of a rotation. This assembly is positioned so that the laser beam passes through the rotational center of the polarizer, as close as possible to the beam expander.

Beam expander with spatial filtering The beam expander was obtained from Oriel Corp., catalog #15260. In this application, this device consists of a 9 mm microscope objective lens focused on to a 25 μm pinhole, followed by a 160 mm collimating lens. This expands the beam 9 times, so that the $1/e^2$ diameter is 5.76 mm.

The purpose of the beam expander is to reduce the angular divergence of the beam, so that the measurements of scattering are truly the scattering from a collimated beam. In the far field, the beam expander results in a

beam divergence of

$$\theta = \frac{\lambda}{\pi w_0} = 0.004^\circ \tag{5.1}$$

where w_0 is the $1/e^2$ radius. Focused through a 50 mm lens, this can be focused to a spot with $1/e^2$ radius

$$w = \frac{f\lambda}{\pi w_0} = 3.5\mu\text{m} \tag{5.2}$$

In this instrument, the lens focuses onto a CCD detector with photodiodes separated by 25 μm. Thus, the angular resolution of the detector, with a 50 mm lens is 0.029°, which is much larger than the beam divergence. Furthermore, the unscattered beam focuses much smaller than the size of the photodetector, so that the unscattered and scattered components are separable.

Solenoid Beam Blocker A solenoid, under computer control, operates a beam blocker. This blocks the beam entirely when requested by the computer, so that back ground light or dark current in the CCD array can be measured and corrected for.

Analyzing Polarizer The effect of the intensity controller is to change the state of polarization through the rotation of the polarizer. For many experiments I have conducted, the effect of polarization changing would be to confuse experimental results, since reflection off surfaces depends on state of polarization, as determined by the Fresnel reflection equations. Thus, a polarizer after the beam expander preferentially passes one polarization. This polarizer must be adjusted with respect to the intensity controller,

which in turn must be adjusted with respect to the polarization of the laser. To produce the minimum intensity, the analyzing polarizer must be *parallel* to the polarization of the laser, and the intensity controller polarizer must be *perpendicular* to the polarization of the laser. The maximum intensity occurs when the laser and both polarizing filters have parallel polarization.

Sample Volume A sample volume must be short enough to prevent vignetting of scattering to be measured. In this work, I use two sample volumes; one is 3/4" thick, the other 1" thick. The volume itself is small enough to allow vigorous stirring by small propeller, driven by a DC motor. The vigorous stirring allows sizing of dense, large particles, including dense spheres. The use of a propeller allows efficient, vigorous stirring without generation of bubbles. Stirring noticeably diminishes the observed random noise in measurements of scattering. The effect of stirring is to reduce the effects of speckle and small number statistics of sparse sizes of particles.

Focusing Lens While one might opt for a custom Fourier transform lens for this application, I use stock camera lenses. In the work reported here, this is a 50mm F1.8 Cannon camera lens. The effective focal length of this lens when measuring the scattering from water samples is 66.5mm, taking into account the index of refraction of water, $\eta = 1.33$. The advantage of camera lenses is low cost, and high quality optics, with multi-layer anti-reflection coatings.

Beam Mask A knife edge painted flat black, and adjusted by a series of translation stages, blocks the unscattered, focused, component of the laser

Product:	EGG RL128G
Number of photodiodes (elements or pixels)	128
Diode center to center spacing	25μm
Diode aperture width	26 μm
Non-uniformity of sensitivity as specified by EGG, typical	$\pm$7 %
Non-uniformity of sensitivity as specified by EGG, maximum	$\pm$10 %
Dynamic Range, defined as Saturation exposure divided by peak to peak dark pattern, with RC100B/104 sample and hold electronics.	100:1 typical

Table 5.1: Specifications of the CCD photodiode array.

beam, while allowing the scattering at larger angles to reach the detector. The unscattered component otherwise saturates the detector.

CCD detector There are a number of options for measuring the intensity at the focal plane of the lens. These include CCD arrays, pyroelectric arrays, and mechanically scanned photodetectors. In this work, I use an EG&G Reticon RL-128G CCD array. Pertinent details of this array are contained in table 5.1. This is a monolithic self-scanning array of photodiodes. Each photodiode in this array is associated with a storage capacitor which integrates photocurrent, and a multiplex switch, for periodic readout. This device is described in detail in EG&G literature [119]. This CCD

device was chosen entirely because of its low cost; other CCD photodiode arrays are available with higher dynamic range, at greater cost.

The CCD photodiode array in this instrument has 128 photodiodes, which I refer to in this work as either "elements" or "pixels." In this instrument, I achieve considerably better dynamic range than the manufacturers specifications suggest, by compensating individual photodiodes for bias (dark current) and gain, and by averaging many scans.

Electronics An IBM personal computer controls the instrument and records data. The other electronic parts of the instrument interface the computer to the stepping motor in the intensity controller, and convert the output of the CCD array for storage and preliminary processing.

Sample and Hold The output of the CCD array is a series of video pulses, which vary in magnitude according the intensity of light on corresponding photodiodes. The series of pulses makes up a *scan.* The electronics supplied with the CCD array provides a clock signal which marks each pulse, as well as timing lines which mark the beginning and end of scans. In this instrument, the video pulses are sampled and held, via a Reticon RC100/104 amplifier, producing an output suitable for analog to digital conversion.

Analog to Digital Conversion The output of the sample and hold amplifier feeds to an 12 bit analog to digital converter; this is a DASH-8 circuit board supplied by Metrabyte corporation, which occupies an expansion slot in an IBM PC. The clock pulses and timing lines coordinate

with the analog to digital converter via some simple CMOS circuitry of our design. Metrabyte supplies subroutines which are callable by the BASIC computer language. I obtained source code in assembly language for these routines, and optimized it to increase the rate of sampling.

The optimized assembly language, in conjunction with the circuity described, sets in to action the following chain of events when a BASIC subroutine requests a scan of data:

1. The subroutine calls an assembly language routine which, through the DASH-8 circuit board, raises a logical line to the CMOS control circuitry of our design, indicating that the computer is ready to accept a digitized scan.

2. The CCD array scans continuously; scans are marked by a *beginning of scan* pulse.

3. The next time a *beginning of scan* pulse is issued, the clock pulses marking video samples are gated to the digital to analog converter. These tell it to make a sample, digitize it and store it in computer memory.

4. When a clock pulse is received in the CMOS circuitry, a logic line is toggled. The computer checks continuously for this line to toggle; when the line toggles, the computer checks for the "sample ready" line of the analog to digital converter to indicate that it has finished digitizing a sample. At this point, the computer reads the digitized sample.

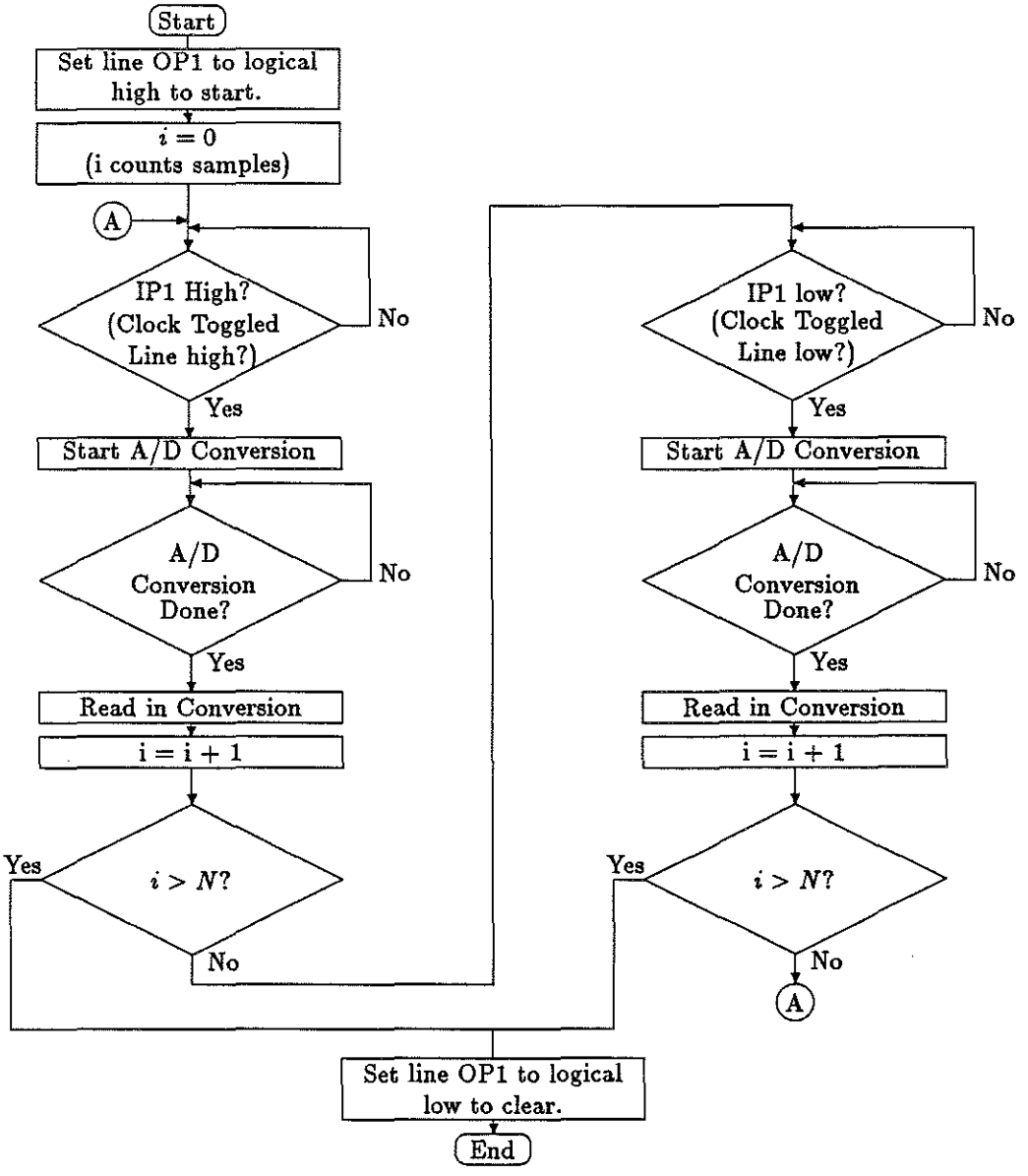

Figure 5.2: Flow chart of the computer code for sampling.

This chain of events is summarized in a flowchart, in figure 5.2.

With the optimized assembly code and additional electronics, the DASH-8 board digitizes data from the CCD array at about 17khz. By contrast, the un-optimized software delivered with the board can digitize data at only 4khz, which is too slow for the CCD array to operate. Since each scan has 128 samples, a scan is read into the computer in 7.5 milliseconds, not including overhead time entering and exiting the subroutine which initiates the sampling.

On the IBM PC for this instrument, I use the Microsoft QuickBasic compiler for all but the assembly language programming. Thus the instrument is under the control of an easy to use, but fast, high level computer language. The computer averages hundreds of scans from the CCD array, obtaining average intensities and statistics of intensity such as standard deviation as well. The final processing, including inversion of the measured intensities and graphical output, is performed on a Digital Equipment Corp. VAX 11/780, but eventually could be performed on the PC as well.

5.2 Suggested Improvements in Instrumentation

Since, as a laboratory prototype, this scattering instrument is intended to address a limited scientific agenda, at low cost, there are a number of improvements that can be made in future instruments, that are immediately evident. I mention these briefly; most concern increasing dynamic range.

There are other photodiode arrays available commercially that have better specifications than the low cost EG&G RL128G array we use. These include, for example, the EG&G CCPD, charge coupled devices, or PCD, plasma coupled devices from other manufacturers. These offer considerably greater dynamic range. For the EG&G CCPD's, dynamic range as determined by saturation exposure divided by fixed pattern noise, is typically 350 versus 100 for the array we use. Of course, this noise is easily compensated for; the important quantity, dynamic range as defined by saturation exposure divided by RMS thermal noise, is typically 5000.

The Fairchild Weston division of Shlumberger, Inc. is presently introducing a series of particularly promising CCD photodiode arrays. One device, for example, offers a typical dynamic range of 7500:1, with 2592 photodiodes with dimensions 10μm x 10μm and 10μm separation. The devices include sample and hold electronics, and additional circuitry to prevent "blooming," or excess spill-over from saturated diodes to other diodes. These devices represent the next generation in CCD detectors. The disadvantage of these devices, actually, is that they provide excessive angular resolution.

Another way to increase dynamic range of the detector is through cooling, since the dark current is halved for each 6.7° reduction in temperature. Cooling can be accomplished easily and cheaply with thermoelectric heat pumps.

Of course, increased dynamic range on the part of the photodetector array is not useful without a similar increase in the analog to digital conversion. I also suggest the use of a 16 bit analog to digital converter.

Data translation makes 16 bit converters for IBM PC's with up to 100,000 samples per second throughput, for example, for approximately $2500. Another possiblility, however, could be to logarithmically amplify the signal before digitizing, and exponentiate after digitizing [120]. In this case it is important to choose a logarithmic amplifier with good *log conformity*, which is analogous to linearity in a linear device. A candidate device is the Burr-Brown LOG100.

Finally I make some suggestions for devices not currently available commercially that would improve this instrument greatly. Rather than using linear photodiode arrays, one could consider arrays consisting of annular rings. Such a detector is used in the Malvern instrument, with 32 rings, but I suggest an array with perhaps 128 rings, for better resolution. Arrays of annular rings have two advantages. First, detector area increases with radius, giving greater light collection at low intensities where it is needed. Secondly, annular rings spatially average the intensity, so that less temporal averaging is needed to reduce speckle.

5.3 Laboratory Procedure

The laboratory procedure for the particle sizing experiment is summarized in figure 5.3. In the following paragraphs I elaborate on each step in the procedure.

The first step is to initialize the instrument, putting it in a state of readiness. This means turning the motorized polarizer to reduce transmitted intensity to a minimum, and moving the beam mask so that the

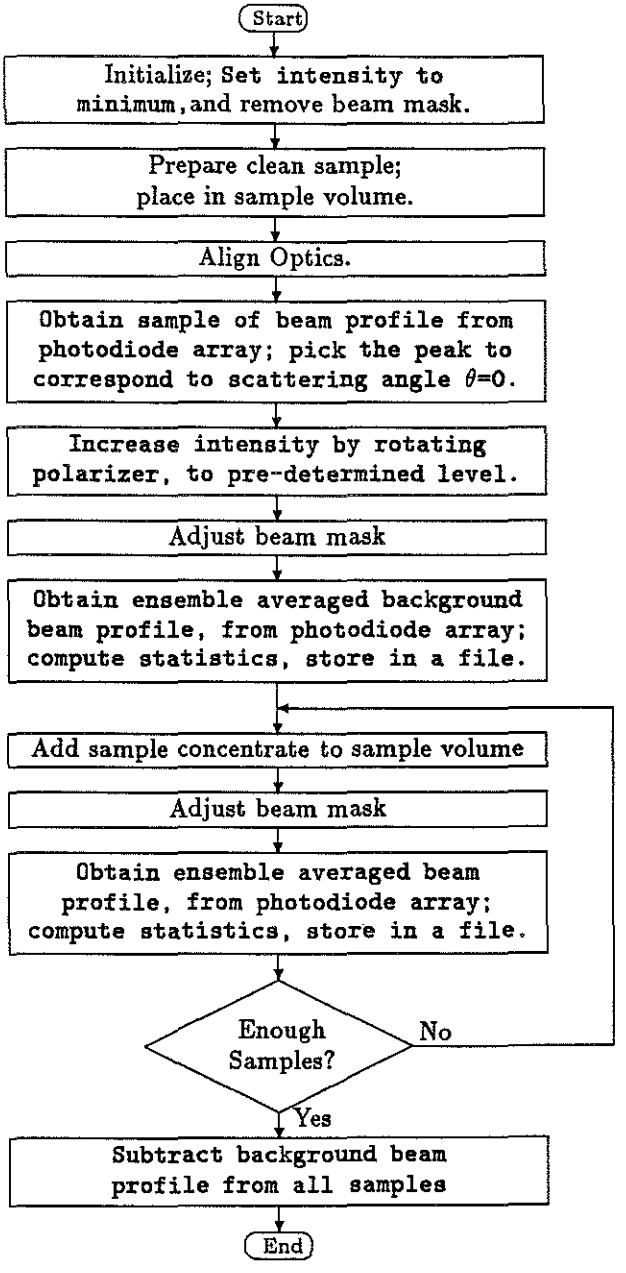

Figure 5.3: Laboratory Procedure for particle sizing. Typewriter font indicates steps performed by computer.

unobstructed beam hits the detector. When this is done, the center of the beam can be observed on an oscilloscope monitoring the output of the sample and hold amplifier. This is used to check optical alignment, and, in a later step, the computer chooses the photodiode registering the most intensity as representing $\theta = 0$.

The next step is to prepare the sample for sizing; this is discussed in section 5.3.1. With the sample placed in the measurement volume, one aligns the optics, checking to see that the photodiode array is positioned so that the maximum intensity is a maximum on one photodiode near one end of the array.

At this stage, the computer takes over most of the experiment. It instructs the motorized polarizer to turn so that the transmitted beam has near maximum intensity. This causes some of the array, especially at small θ, to saturate. The saturation is then eliminated by manually moving the beam mask into place, to eliminate all effects of saturation, as monitored on an oscilloscope.

The next task is to obtain a background sample of intensity, to be subtracted from all subsequent measurements. This, and all subsequent measurements, are averages of many scans of the photodiode array. The standard deviations of these measurements from the mean are also computed. The numerical processing of data as it is collected is discussed in detail in section 5.4.1.

With the background scattering thus measured, the instrument is ready to measure scattering from samples. One or more drops of solution containing the sample of interest are added to the sample volume. The beam mask,

is then readjusted, if required, to eliminate any saturation. The computer then averages scans from the photodiode array, as in the measurement of the background. Usually this process repeats several times, each time with one or more additional drops of concentrated solution of particles.

Finally, when all data is gathered, the computer subtracts the background scattering from each set of measurements. Since some photodiodes at small θ are saturated, some small angle data is discarded at this stage. The data is then ready for inversion.

5.3.1 Preparing Samples for Particle Sizing

Since this work involves a laboratory prototype, we have a luxury not available with an *in situ*, or field, instrument — one can prepare the samples analyzed. As long as the appropriate details of the preparation of our samples are recorded, the data can still be used to design instruments to analyze samples in the field.

The advantage of samples prepared in the lab are several. First, one can prepare samples with abnormally high concentrations of one particle size. This results in stronger, more easily measured scattering, and reduces the effect of unwanted foreign particles. One need not go to extreme measures to obtain ultra-pure water for "background" samples. Furthermore, this allows tests with special distributions of test particles. These include for example, narrow Gaussian distributions of polystyrene spheres, which are discussed in further detail in later sections.

One can also increase concentrations of particles between sets of mea-

surements. This provides a means to determine whether or not multiple scattering is a problem, as discussed in section 2.4.

In addition to preparing samples specially, a motor stirs the sample volume during scattering measurements. This, in conjunction with temporal averaging, reduces two sources of noise. These are speckle noise, discussed in section 2.7, and the small number statistics of sparsely populated sizes.

5.4 Inversion of the Data: Statistical Issues

5.4.1 Preprocessing of Experimental Data

In addition to averaging sets of scans of small angle scattering data, the computer also computes a standard deviation from the mean for each angular sample (photodiode). In principle, this information can be incorporated into inversions, so that the noisiest data is weighted least. In practice, these statistics provide a measure of the quality of the data. For example, in the experimental data I present in the next chapter, there is little random noise. From this, I conclude that for this data, speckle noise, and small number statistics of scattering from scarce species, are insignificant.

My experimental procedure is to average K *sets* of *scans* of the CCD array. Each *set* is actually an average of 10 *scans* from the array. The standard deviation we compute is based on the averaging of the scans.

I will summarize the data collection process in mathematical form as follows. Let S_{ij} represent the digitized output from the CCD array from the i-th element, on the j-th scan, with the individual dark background biases

subtracted from each element (photodiode). Then, the average intensity observed is calculated as

$$I_i^{\text{obs}} = \frac{1}{K}\sum_{k=1}^{K}\left(\frac{1}{10}\sum_{j=1}^{10} S_{ij}\right) \tag{5.3}$$

and the variance is computed as

$$(\sigma_i^{\text{obs}})^2 = \frac{1}{K}\sum_{k=1}^{K}\left(\frac{1}{10}\sum_{j=1}^{10} S_{ij}\right)^2 - \frac{1}{K}\left(\sum_{k=1}^{K}\frac{1}{10}\sum_{j=1}^{10} S_{ij}\right)^2 \tag{5.4}$$

The variances are computed over sets of scans in this fashion to decrease the time for computation, because this is not post-processing — it is performed during an experiment.

From all sets of measurements of particle distributions, a background is subtracted, which is also an average of K sets of 10 scans, for which a standard deviation has been computed. The scattered intensity is thus

$$I_i^{\text{sca}} = I_i^{\text{obs}} - I_i^{\text{background}} \tag{5.5}$$

Assuming Gaussian, uncorrelated, independent noise, one can compute a variance

$$\sigma_i^2 = (\sigma_i^{\text{obs}})^2 + (\sigma_i^{\text{background}})^2 \tag{5.6}$$

For some photodiodes, where the intensity is very small, I_i^{sca} is negative. For these diodes, σ_i^2 is re-assigned an arbitrary but large value.

At this stage the computer also corrects I_i^{sca} for the relative gains of individual photodiodes, which were measured using diffuse (incoherent) white light. For the CCD array in these experiments, the relative gain varies by as much as $\pm 4\%$ before correction.

Typically I choose K, the number of sets of scans, to equal 40. Thus, 400 scans are averaged. This amount of averaging reduces Gaussian noise by 13db.[1] We chose this number of scans somewhat arbitrarily; it is apparently more than needed, but they take only 1 minute to complete.

5.4.2 Statistics of Inversions: Sources of Noise

There are a number sources of noise which are potentially of concern. Two of these, the dark pattern, the pattern the array "sees" in the dark, and non-uniformity of gain, we attempt to compensate for. Noise sources specific to CCD photodiode arrays are discussed in references [119], [121] and [122]. The noise sources we cannot compensate for, except by averaging, I discuss in the following paragraphs.

Random, Gaussian Noise One class of noise, which can be eliminated by averaging, is random, Gaussian noise; this has contributions from small number statistics of sparsely distributed particles, from speckle, from thermal (Johnson) noise, and dark current noise in the photodiode array. Of these processes, those that are not Gaussian, contribute Gaussian noise in

[1]Since so many scans are averaged, random noise is probably *much* smaller than systematic errors or *bias*. This is one reason why I do not compute confidence intervals, or error bars, for the results of our inversions. Were random noise dominant, one could compute a covariance matrix $\mathbf{C}$ for the estimate $\hat{\mathbf{m}}$ from

$$\mathbf{C} = \mathbf{H}\mathbf{C}_\mathbf{d}\mathbf{H}^\mathsf{T} \tag{5.7}$$

where $\mathbf{C}_\mathbf{d}$ is the covariance matrix of the data, and is the "inverse" of $\mathbf{G}$. To do this, however, would be extremely misleading. Another reason is that many algorithms that find the estimate $\hat{\mathbf{m}}$ do not explicitly compute an inverse matrix $\mathbf{H}$. These include, for example, the iterative algorithms.

the averaging process according to the central limit theorem.

Systematic, Correlated Noise A second set of sources of noise is stray light, unwanted reflections, blooming (signal cross talk between adjacent photodiodes, due to charge spill over). One cannot compensate for these, but can reduce the effects by minimizing the number of optical elements, and using coated lenses.

Quantization Noise A third noise is due to the digitization or quantization by analog to digital converter, which is evident in some of our data at low intensities.

Nonlinearity Another noise contribution is nonlinearity of the CCD detector; there are no manufacturers specifications for this, and I assume that it is small.

Drift The laser power fluctuates, so the background intensity measured early in an experiment may not be appropriate for scattering measured later. Over an eight hour period, after 15 minutes of warmup, according to the manufacturers specifications, power may vary as much as $\pm 5\%$. In our work, an entire experiment usually lasts less than 30 minutes.

5.4.3 Statistics of Inversions: Maximum Likelihod Estimates

The statistics of measured intensity are the consequence of a variety of processes. In inversions, one would, in principle, like to weight the least noisey measurements to get the best estimate of size distribution, $n(x)$. For some data, a useful approach may be to assume that noise from each photodiodes is Gaussian, with the mean and variance described in section 5.4.1. While this is not accurate where quantization noise is most severe, or where other sources of noise dominate, as is true for the results in this thesis, it may be true in other work.

By assuming Gaussian noise, measurements of scattering have a probability density of the form

$$P_{\mathbf{I}^{\text{meas}}}(\mathbf{I}^{\text{meas}}) = \frac{1}{\sqrt{(2\pi)^K \det\Lambda}} \exp\left((\mathbf{I}-\mathbf{I}^{\text{meas}})^{\mathsf{T}}\Lambda^{-1}(\mathbf{I}-\mathbf{I}^{\text{meas}})/2\right) \qquad (5.8)$$

where $\mathbf{I}$ is a vector of scattered intensities, and $\mathbf{I}^{\text{meas}}$ is a vector of the measured intensities. The matrix Λ is the covariance matrix, given as the expectation

$$\Lambda = \mathrm{E}\left[(\mathbf{I}-\mathbf{I}^{\text{meas}})(\mathbf{I}-\mathbf{I}^{\text{meas}})^{\mathsf{T}}\right] \qquad (5.9)$$

The matrix equation

$$\mathbf{I} = \mathbf{T}\mathbf{n} \qquad (5.10)$$

represents the scattered intensity $\mathbf{I}$. We seek an estimate $\hat{\mathbf{n}}$ to $\mathbf{n}$. A reasonable approach is to determine $\hat{\mathbf{n}}$ to be the size distribution which maximizes equation (5.8), the probability of the measured scattered intensity. To do

this, the quantity

$$(\mathbf{I} - \mathbf{I}^{\text{meas}})^{\mathsf{T}} \Lambda^{-1} (\mathbf{I} - \mathbf{I}^{\text{meas}}) \tag{5.11}$$

should be minimized. This is a weighted least square norm, and the solution, $\hat{\mathbf{n}}$ is known as the *maximum likelihood solution.*

Assuming the noise in individual photodiodes to be statistically independent and uncorrelated, the matrix Λ reduces to individual variances, σ_i. If we define new vectors of intensity

$$[I^{\text{meas}}]'_i = \frac{[I^{\text{meas}}]_i}{\sigma_i} \tag{5.12}$$

and a new matrix

$$[T]'_{ij} = \frac{[T]_{ij}}{\sigma_i} \tag{5.13}$$

then (5.11) can be written

$$(\mathbf{T}'\mathbf{n} - \mathbf{I}^{\text{meas}\prime})^{\mathsf{T}} \Lambda^{-1} (\mathbf{T}'\mathbf{n} - \mathbf{I}^{\text{meas}\prime})^{\mathsf{T}} \tag{5.14}$$

This reveals that the maximum likelihood solution $\hat{\mathbf{n}}$ can be obtained simply by dividing the intensities and the corresponding rows in the matrix $\mathbf{T}$ by the standard deviations at respective pixels.

In a previous section I introduced the rephrased inverse problem, so that we seek $m(x) = xn(x)$ from $d(\theta) = \theta^3 I(\theta)$. To determine the maximum likelihood estimate for $m(x)$, replace the data vector $\mathbf{d}$ by a vector d', given by

$$d'_i = \frac{d_i}{\sigma_i} \tag{5.15}$$

and replace the matrix $\mathbf{G}$ with a new matrix $\mathbf{G}'$

$$G'_{ij} = \frac{G_{ij}}{\sigma_i} \tag{5.16}$$

In practice, it may not be desirable to seek the Maximum likelihood estimate size distribution, because the matrix error-weighted $\mathbf{G}'$ is more singular than the non-weighted matrix $\mathbf{G}$. Iterative algorithms may take much longer to converge, if they converge at all. Furthermore, at points of low scattered intensity, noise is usually high - a Maximum likelihood algorithm may ignore points of low intensity altogether. This is particularly serious for narrow size distributions, for which the scattering has troughs, which are usually noisy. Thus, in this work, I do not present the results of Maximum likelihood inversions, because they offer no improvements for the inversions shown.

Chapter 6

Inversions of Experimental Small Angle Scattering Data

In light of the theoretical development of chapters 3 and 4, inversion of small angle data is most robust when $d(\theta)$ is sampled evenly in the interval $(0, \theta_u]$.

In practice, however, this condition may not be met. For example, for very narrow distributions, $d(\theta)$ may approach the asymptote slowly, if at all, within the range of measurements. In chapter 3, I used the Heisenberg uncertainty principle, to derive the *lower bound*

$$\theta_u > \frac{1}{4x_u} \tag{6.1}$$

where θ_u is an angle at which $d(\theta)$ has reached, within experimental error, its asymptote. Recall that x_u is the largest size in a distribution, or more generally, the "width" of a size distribution. Thus, for a narrow size distribution, for which x_u is small, θ_u is large, perhaps out of the range of measurements. The function $d(\theta)$ can also approach its asymptote slowly

for a distribution with a large x_u, if the distribution has high frequency content.

Data may also be missing at the smallest angles. In experimental measurements, limited dynamic range and the inability to mask out the unscattered beam may each invalidate data at the smallest angles. In our experimental work, for example, the first 10 to 20 diodes must be masked to avoid these problems. In the worst case, $d(\theta)$ approaches its asymptote very rapidly, so as to be indistinguishable, even at the smallest angle measurement. The data measured thus consists only of the asymptote of $d(\theta)$.

Measurements with these deficiencies present special problems for inversion. Where $d(\theta)$ approaches its asymptote slowly, for example, $n(x)$ has high frequency content; that is, it has sharp peaks, or is, perhaps, discontinuous. Sharp peaks present a special problem for inversion. Recall that the matrix equation $\mathbf{d} = \mathbf{Gm}$ approximates an integral equation, with good approximation if the integration increment Δx is small enough. The integral equation has two parts, approximately; a Fourier sine component, and a constant component. Recall that the matrix $\mathbf{G}$ is approximately

$$\mathbf{G}_{ij} \approx \frac{\Delta x}{\pi k^3}(1 - \sin 2x\theta) \tag{6.2}$$

If $m(x)$ is appropriately band-limited, the integration increment Δx is small enough that the Fourier component represents the related integral $\int \sin 2x\theta$ accurately. This is not the case for the constant component, however. Even if a size distribution $m(x)$ is band-limited, it must be smooth enough so

that the approximation

$$\sum_i m(i\Delta x)\Delta x \approx \int_0^{x_u} m(x)\,dx \tag{6.3}$$

is good. Otherwise, the matrix equation $\mathbf{d} = \mathbf{Gm}$ is not an accurate approximation to the integral equation it replaces. When $\mathbf{G}$ is used to invert experimental data, the experimental data may appear infeasible to an inversion algorithm, because $\mathbf{d} = \mathbf{Gm}$ is a poor approximation to physical reality.

In addition, when θ_u is well out of the range of measurements, so that it is difficult to estimate, it may not be possible to insure that the Nyquist criterion is satisfied. For such measurements, the best solution may be to introduce *a priori* information by placing bounds in x on $n(x)$. The effect of placing bounds $x_{min} > 0$ and x_{max} is to change the Nyquist criterion,

$$\Delta\theta \leq \frac{\pi(M-1)}{2x_u} \tag{6.4}$$

of section 3.2.5 to

$$\Delta\theta \leq \frac{\pi(M-1)}{2\Delta x} \tag{6.5}$$

where the upper bound x_u has been replaced by the width $\Delta x = x_{max} - x_{min}$. Since $\Delta x < x_u$, the necessary sampling interval, $\Delta\theta$ is correspondingly larger. I present examples of examples of inversions of experimental data with bounds imposed on x in section 6.3.

Missing data at small angles results in instability in inversions; we saw this in section 3.2.6.

Sampling deficiencies such as these may be mitigated, by robust algorithms, or *a priori* information regarding $n(x)$. Algorithms can be made

more robust by applying constraints such as positivity smoothness to $n(x)$. In the following sections, I examine the results of inversions experimental scattering data from size distributions prepared in the laboratory. First, I use synthetic data to illustrate that inversions are unstable for uneven sampling. Next I compare inversion algorithms; then, I demonstrate introduction of *a priori* information.

6.1 Inversion of Synthetic Data

I have stated that inversions of experimental data are most robust when the data is evenly spaced between 0 and θ_u, and are unstable when the data is unevenly spaced, or is offset in θ. This is most easily seen with synthetic data, since the prototype instrument can only collect data which is offset in θ.

To illustrate the instability resulting when measurements of scattering are offset in θ, I have synthesized and inverted two sets of measurements, resulting from the same size distribution. The size distribution $n(x)$ is Gaussian, with mean $x = 200$ and standard deviation 5.5%, or $\sigma_x = 11$. I chose this distribution because it is similar to one I measured in experiments, described in the next section.

The intensity from this distribution was constructed as a sum over 256 points in x, from $x = 150$ to 250, using diffraction theory. In the first set of synthetic data, I constructed $I(\theta)$ at 95 points, from $\theta = 0.0218°$ to $\theta = 2.04°$. For this set, only the data from the photodiode at $\theta = 0$ is missing. For the second set of synthetic data, which is similar to the

12dp400, I constructed $I(\theta)$ the same way, except that the data are offset in θ by 22 photodiodes. These are plotted as $d(\theta)$ in figure 6.1.

I present the results of inversions of these distributions in figure 6.2. The inversions are based on diffraction theory as a forward model. Both employ the Phillips-Twomey algorithm, discussed in section 4.2.2, and represent inversions of $d(\theta)$ to yield $m(x)$, which was divided by x to yield $n(x)$.

For both sets of measurements, inversions were performed only for 30 sizes. The matrices representing the scattering were constructed according to the recommendations of section 4.1.2, with $\theta_u = \theta_{max}$, but truncated so that only the 30 smallest sizes are represented. The result is an over-determined matrix, with 95 points of data and 30 unknowns. The principle reason for this was to reduce computing time, but the effect is to introduce *a priori* knowledge regarding the largest particles in the distribution. This increases the stability of the inversions as well.

The damping parameter in each inversion was $\nu = 100$. I chose this value empirically; it yields nearly optimal results.

The inversions in figure 6.2 demonstate the instability of inverting measurements where data is missing, or where the data is offset in θ. Both inversions plotted in figure 6.2 should indicate a single peak at $x = 200$, representing the Gaussian size distribution from which the scattering data was synthesized. The lower curve, for which the scattering measurements were offset by one photodiode, shows such a peak, with some side lobes. The upper curve, however, which represents the inversion results of the data offset by 22 photodiodes, does not resemble a Gaussian at all.

Though θ_u was chosen equal to θ_{max}, when constructing the matrix

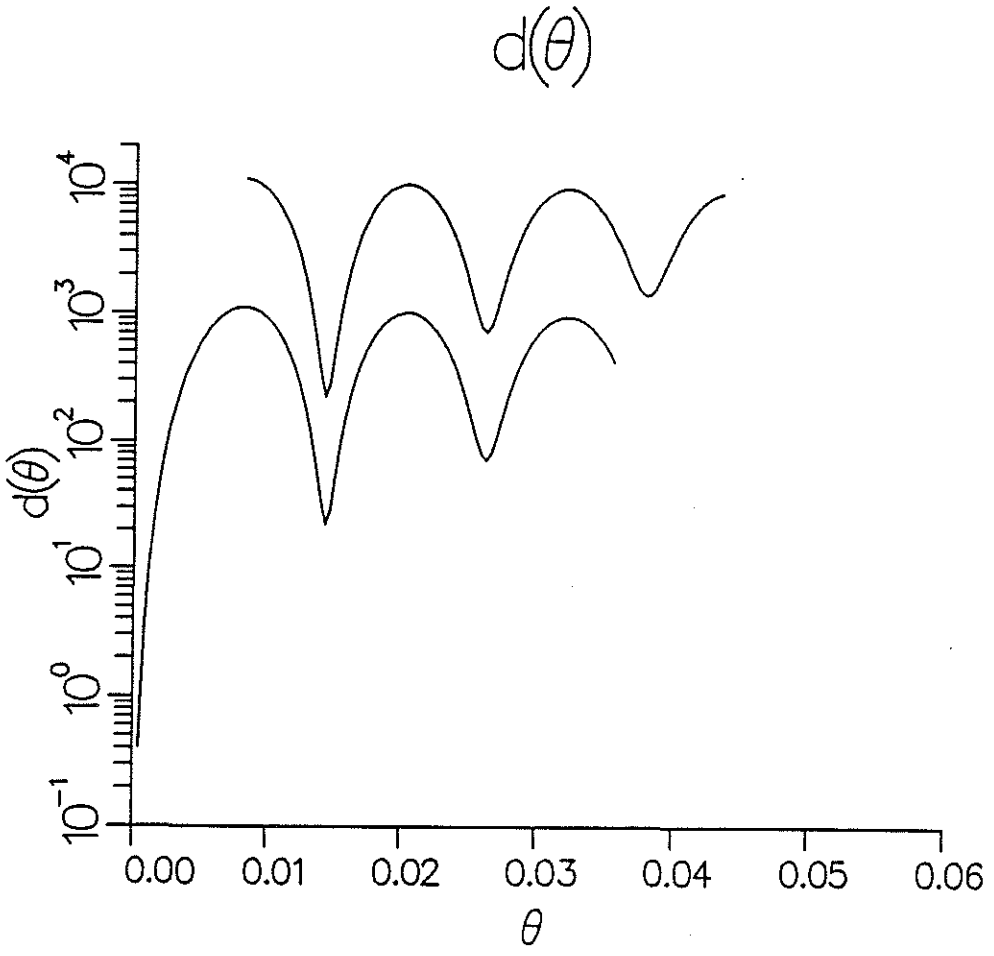

Figure 6.1: $d(\theta)$ functions, synthesized for Gaussian distributions, with center at $x = 200$ and standard deviation 5.5%. The upper plot represents the second set of data, and is offset vertically for clarity. The size distribution was chosen to be similar to that of the sample 12dp400, discussed in later sections. The angles of the measurements in the upper plot correspond to the angles at which scattering was observerd for the sample 12dp400.

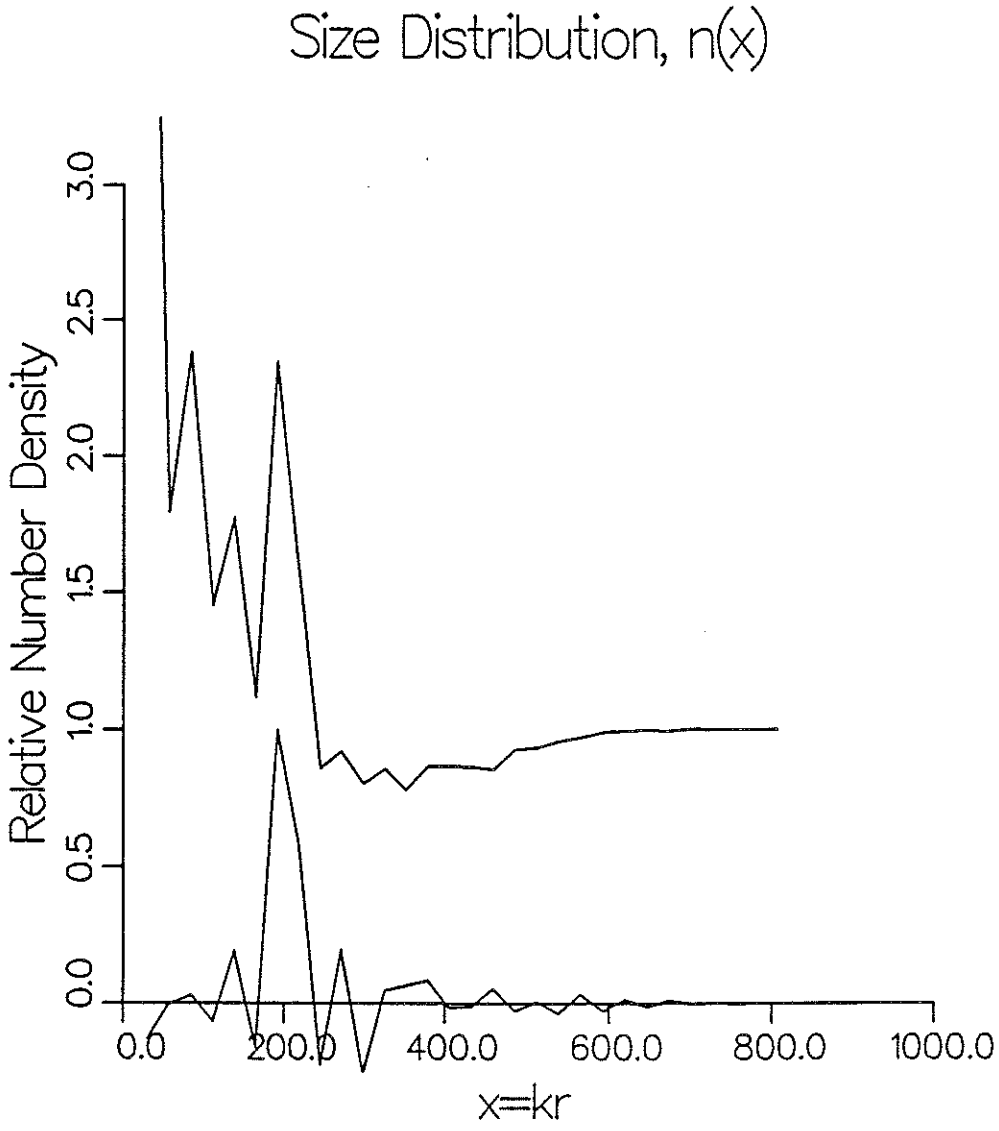

Figure 6.2: Results of inversions of the synthetic data. The upper plot represents the inversion of the upper set of data in figure 6.1, and is offset vertically, for clarity.

representing forward scattering, clearly θ_u is larger than θ_{max} the case for this data. This suggests that $m(x)$ has high frequency content, and raises a second interesting point, regarding the side lobes in the lower curve in figure 6.2. These are purely the result of representing the scattering by a matrix G in which the integration increment Δx is too small. Although the Nyquist criterion was satisfied, the requirement that $\int m(x)\,dx$ approximate $\sum m(i\Delta x)\Delta x$ was not. Thus, for this data, $\mathbf{d} = \mathbf{Gm}$ is a poor approximation to the integral it approximates.[1] In a subsequent section, I demonstrate inversions of similar data, which incorporate *a priori* information to eliminate these side lobes.

6.2 Comparison of Inversion Algorithms

In a previous section I reviewed some inversion algorithms from a theoretical stand. Here, I compare the performance of these algorithms in inverting real data. These algorithms are also readily compared with synthetic data; these results, in our experience, would not differ significantly. I discussed the collection of this data in section 5.1.

This comparison considers the following algorithms:

1. A singular value decomposition method, in which eigenvectors corresponding to small eigenvalues are discarded. See section 4.2.1 for a description.

[1] For a different, smoother size distribution, however, $\mathbf{d} = \mathbf{Gm}$ may be a good approximation; then there would be no side lobes.

2. The Phillips-Twomey Algorithm, of section 4.2.2.

3. Landweber iteration, with and a positivity constraint, of section 4.2.3.

4. Chahine iteration, with a positivity constraint, of section 4.2.3.

5. A "Projection" iteration, with and without a positivity constraint, of section 4.2.3.

6. A non-negative least squares algorithm, (NNLS), of Lawson and Hanson, of section 4.2.3.

7. The Chin-Sliepcevich integral transform, derived in appendix A.

These algorithms fall roughly into two classes; those which do not constrain the distribution $n(x)$ to be positive, and those which do. In the next subsections I discuss the performance of these algorithms, according to the effects of missing data, different size distributions, and noise.

The Chin-Sliepcevich integral transform algorithm is presented here for comparison with the other algorithms. This algorithm is an evaluation of the integral

$$n(x) = \frac{-2\pi k^3}{x^2} \int_0^\infty J_1(x\theta) Y_1(x\theta) x\theta \frac{d}{d\theta} \left[\theta^3 I(\theta)\right] d\theta, \qquad (6.6)$$

which is derived in appendix A. In this work, the derivative in the integral is performed as a central difference. Agrawal and Riley describe a more elaborate algorithm, in which the derivative is performed with an optimal Parks-McClellan filter [123].

Computation of Optical Depth, τ and Attenuation, α. Along with some of the inversions of small angle scattering from polystyrene spheres, I have also calculated the optical depth, τ, for the sample analyzed, and the attenuation coefficient α. The optical depth, τ, is useful for determining whether or not single scattering is a valid approximation. The attenuation coefficient α is useful for relating these results to measurements in the ocean.

In section 2.4 we saw that single scattering is a good assumption when the optical depth τ is small, that is, $\tau < 0.1$. Recall that τ is defined in terms of the intensity of a beam being reduced by extinction to $e^{-\tau}$ of its original value.

The polystyrene spheres generally have a Gaussian distribution, and are supplied suspended in solution. The manufacturer specifies for each vial of spheres, a mean diameter $\overline{x}$, standard deviation, and number density of particles per milliliter. The standard deviation is small, typically 5 percent or less, so, for estimating optical depth, it is safe to assume that all particles have the same size as the mean.

We introduce these spheres to our sample volume, which has cross-sectional area A. The spheres have a total cross-sectional area $n\pi\overline{x}^2/4$, where n is the number of spheres introduced. We compute n from the number of drops added, so n is approximate. I assume the spheres to have an extinction coefficient, Q_{ext} equal to 2, since the extinction coefficient is, in general, asymptotic to 2 as $x \to \infty$. This is evident in figure 6.3, where the extinction coefficient is plotted versus particle radius. This assumption is quite good for particles larger than 4μm. The power loss of a beam

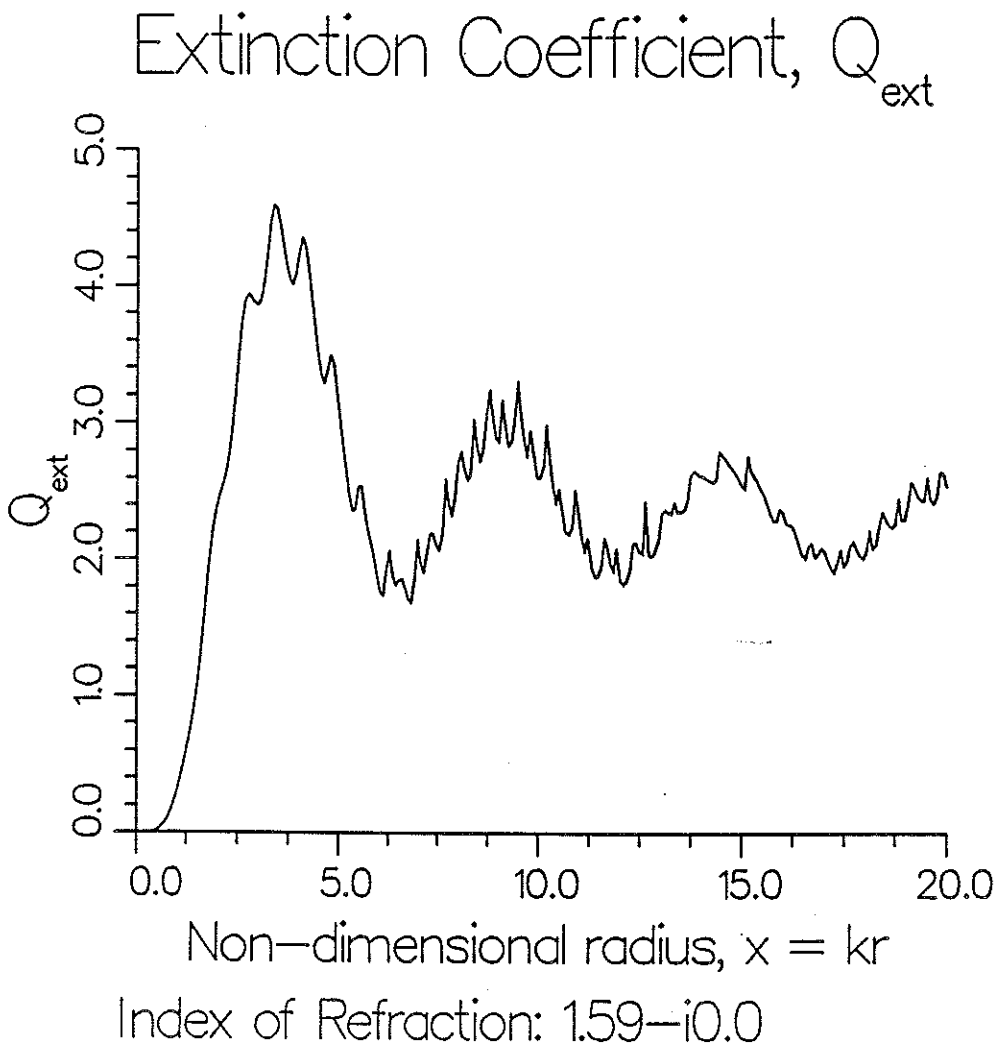

Figure 6.3: The extinction coefficient for the polystyrene spheres, assuming an index of refraction η = 1.59-0.0i.

passing through the sample volume is

$$e^{-\tau} = 1 - n\frac{\pi d^2/4}{A}Q_{ext} \approx 1 - \frac{n\pi d^2}{2A}. \tag{6.7}$$

Thus τ is

$$\tau = -\log\left(1 - n\frac{\pi d^2/4Q_{ext}}{A}\right) \approx -\log\left(1 - \frac{n\pi d^2}{2A}\right) \tag{6.8}$$

The attenuation coefficient is $\alpha = \tau/L$, where L is the actual depth of the sample.

6.2.1 Missing Data

Because, as I have shown, most of the information regarding size distribution is contained in the smallest angles, the effect of missing data at small angles is to degrade inversions. From examination of inversion of experimental data, it appears that the primary effect of missing data at small angles is that *negative* number densities become feasible. With missing data, I find that inversion algorithms which do not constrain $n(x)$ to be positive tend to find meaningless size distributions dominated by negative concentrations. On the other hand, those that do constrain $n(x)$ to be positive yield reasonable results, with some exceptions. As examples, I next compare the inversion algorithms for two sets of data, which differ primarily in the size of the gap in data at small angles.

In the next series of figures, I draw upon scattering data taken from a solution containing a Gaussian distribution of polystyrene spheres, with mean diameter 40.3μm, or non-dimensional radius $x = 200.15$, with a standard deviation of 5.5%. This distribution is of interest because there is

Name of measurements:	12dp400
Type of Particles:	Polystyrene Spheres
Distribution:	Gaussian
Mean Diameter:	40.3μm
Standard Deviation:	5.5%
Source:	Duke Scientific, Catalog # 243.
Concentration in Sample Volume:	5700 ml^{-1}
Estimated Optical Depth, τ	0.127
Attenuation Coeff., α	5.0 m^{-1}

Table 6.1: Relevant data regarding the distribution 12dp400.

a significant gap in the data at small angles, and, in addition, the function $d(\theta) = \theta^3 I(\theta)$ shows few signs of approaching an asymptote with in the range of measurements. I present results from this distribution not because they are typical, but because the distribution illustrates the challenges posed by missing data.

For reference, this distribution is named 12dp400. Relevant data regarding this distribution are shown in table 6.1. The concentration of particles was estimated by counting the number of drops of concentrated stock solution added to the sample volume, for which the number density of particles was determined by Duke Scientific, the supplier. The concentration for this sample is particularly high, resulting in an optical depth $\tau = .12$, but the observed scattering compares well with scattering for lesser concentrations, so multiple scattering is probably not important.

The scattering measured from this distribution is shown in figure 6.4; $d(\theta)$ is shown in figure 6.5. The measured noise associated with these measurements is shown in figure 6.6. Relevant data regarding the measurement of scattering from this distribution are presented in table 6.2.

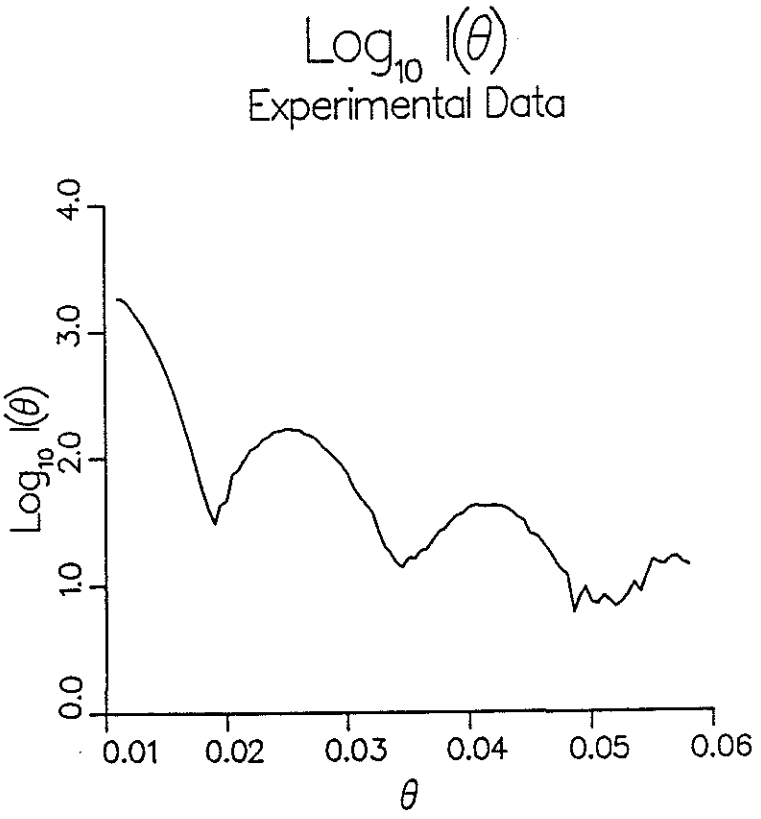

Figure 6.4: Scattering measured from the distribution 12dp400.

Name of measurements:	12dp400	
Number of measurements	95	
Minimum Angle	0.0083	radians
Maximum Angle	0.0435	radians
Number of missing measurements at small angles	21	pixels
Number of CCD scans averaged	400	

Table 6.2: Relevant data regarding the measurement of the distribution 12dp400.

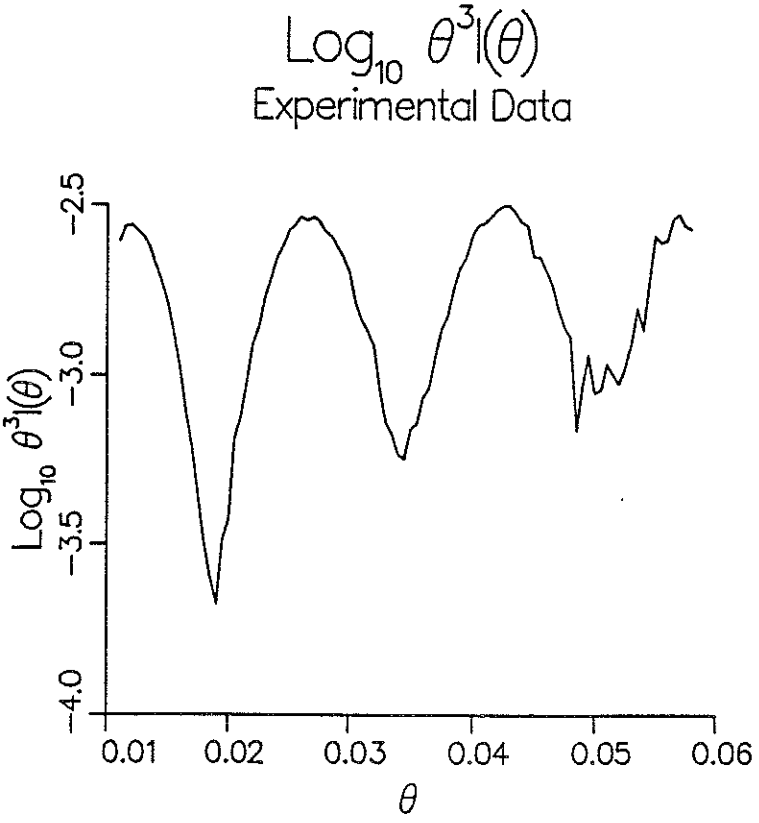

Figure 6.5: Scattering measured from the distribution 12dp400, presented as the function $\theta^3 I(\theta)$.

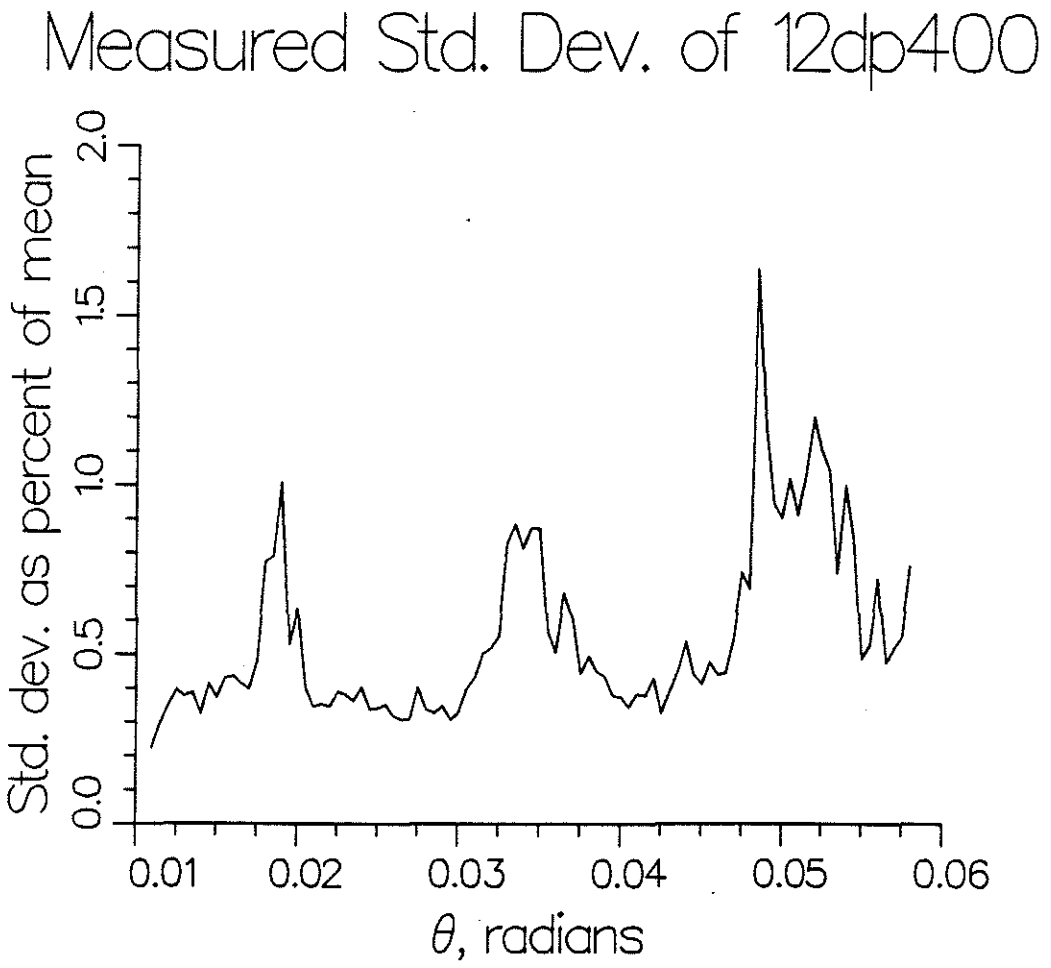

Figure 6.6: Standard deviation of the noise of the scattering measured from the sample 12dp400. This is plotted as a percentage of the mean. The noise is highest at nulls in the intensity.

Name of measurements:	12dp400
Minimum x inverted for	30.9
Maximum x inverted for	807.0
Number of sizes inverted for	30
Forward Model for scattering	Diffraction theory

Table 6.3: Information common to all inversions of 12dp400

Information common to all the inversions is presented in table 6.3.

While 95 measurements were obtained, I performed inversions for 30 sizes only, as I did for the synthetic data in section 6.1. The matrix representing the scattering was constructed according to the recommendations of section 4.1.2, with $\theta_u = \theta_{max}$, but truncated so that only the 30 smallest sizes are represented. The principle reason for this was to reduce computing time, but the effect is to introduce *a priori* knowledge that no particles larger than $x = 807.0$ are to be found in the distribution. Thus, this knowledge increases the stability of the inversions.

Though the results of all inversions are presented as the number density $n(x)$, all inversions in this section in fact, represent inversion of $d(\theta) = \theta^3 I(\theta)$ to determine $m(x) = xn(x)$. The results of the different inversion algorithms for this set of measurements are shown in figure 6.7, with a legend and summary in table 6.4. Each result should show a sharp peak at $x = 200$, corresponding to a narrow Gaussian distribution. The resolution of the inversions is comparable to the known standard deviation of the distribution, so no conclusions regarding the width of the peak can be made from these inversions. The algorithm with best results is the Chahine algorithm, followed by projection iteration. The NNLS algorithm apparently

yields good results for this distribution, but this is not true in general for other distributions.

The feature of this data I want to emphasize is that 21 pixels at the smallest angles were discarded due to saturation of the detector. Only the algorithms which impose positivity on $n(x)$ produce plausible results; these are all iterative algorithms. Because the convergence is not perfect, I present figure 6.8, comparing the functions $\theta^3 I(\theta)$ measured and predicted for the Chahine algorithm.

For some of the inversion algorithms, I have computed a residual, of the form

$$R = \sum_{i=1}^{M} \left(\frac{I_i^{\text{meas}} - I_i^{\text{calc}}}{I_i^{\text{meas}}} \right)^2 \quad \text{for } I_i^{\text{meas}} \neq 0. \tag{6.9}$$

Here I_i^{meas} represents the measured scattering, at pixel i, and I_i^{calc} represents the scattering computed from the estimated size distribution $n(x)$. A residual of this form was chosen because it measures fit at low intensities with the same weight as the fit for high intensities. This residual is reported for many of the results. It is worth noting that although the Phillips-Twomey algorithm produces negative number densities, it often produces very small residuals, corresponding to near perfect fits.

Concluding, missing data at small angles make inversion algorithms susceptible to negative number densities unless $n(x)$ is constrained to be positive. In my experience, it is important to have as much data as possible at small angles, even if the price is increased noise at large angles where the intensity is small.

In general, as for this data, three algorithms appear to produce the

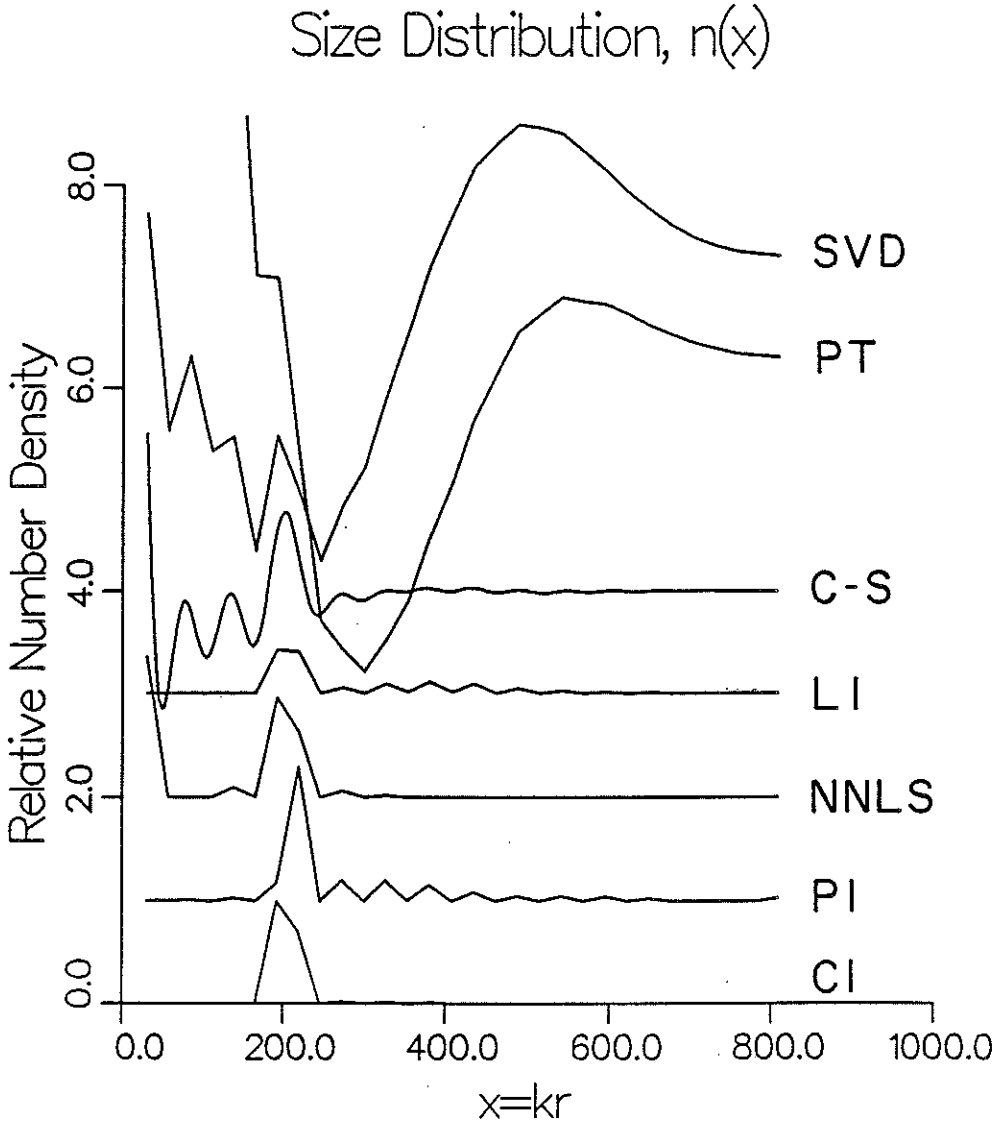

Figure 6.7: Comparison of inversion algorithms on 12dp400. Curves are offset, appearing from top to bottom in the same order as in table 6.4.

Algorithm	Positive Cons-traint?	Comments	Residual
SVD (SVD)	No	Algorithm Fails (Negative Densities) 1 eigenvector discarded.	
Phillips-Twomey (PT)	No	Algorithm Fails (Negative Densities) $\nu = 1$	
Chin - Sliepceich Integral transform (C-S)	No	$n(x)$ found at 256 points.	
Landweber Iteration (LI)	Yes	Peak at $x = 200$, but additional peaks as well. These might have disappeared with further iteration; Convergence is slow.	0.80
NNLS (NNLS)	Yes	Fast Convergence but unreliable results	0.33
Projection Iteration (PI)	Yes	Strong Peak at $x = 200$ A Plausible Result	1.9
Chahine Iteration (CI)	Yes	Strong Peak at $x = 200$ Perhaps the best result. Fast Convergence.	0.20

Table 6.4: Summary, comparison of inversion algorithms on 12dp400 data, which is missing 21 points of small angle data. Results appear from top to bottom in the same order as in figure 6.7. Curve labels appear in parentheses.

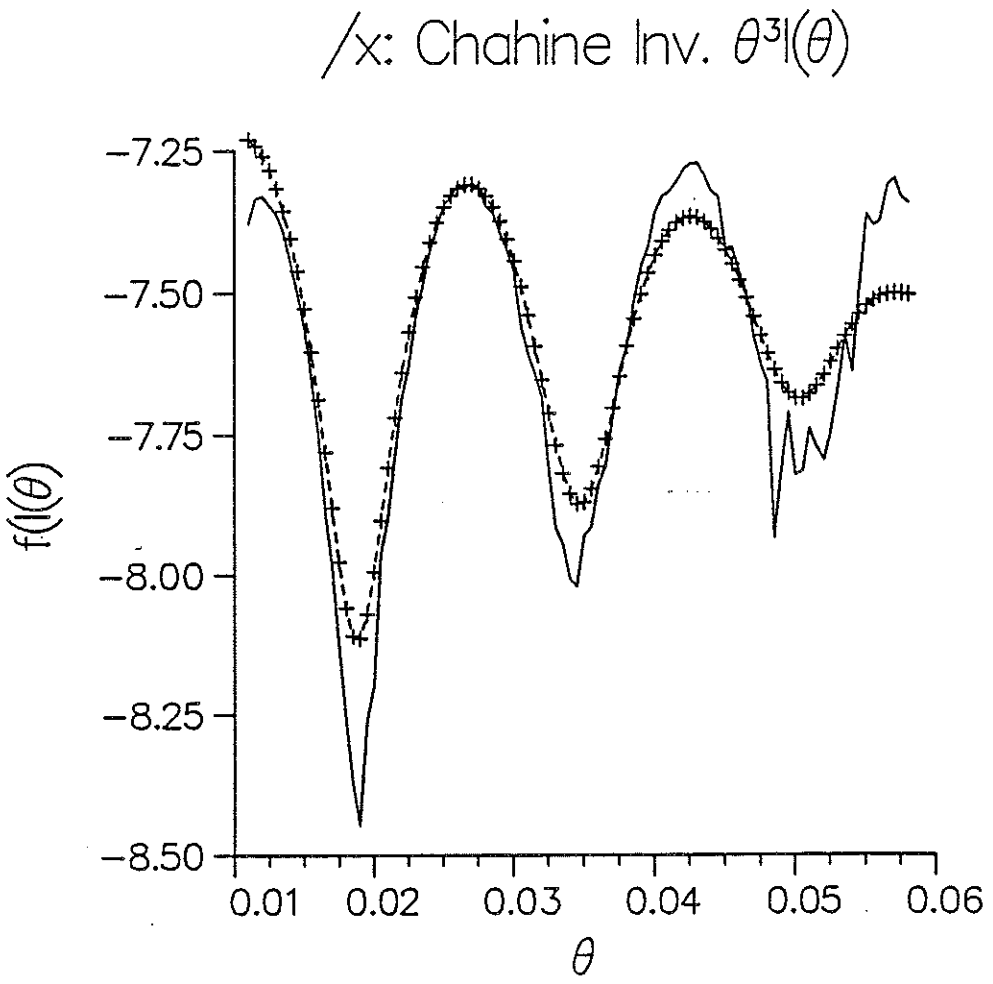

Figure 6.8: $\theta^3 I(\theta)$ measured (solid line) vs. predicted from computed $n(x)$ (crosses) for Chahine algorithm. This shows the quality of fit when iteration was terminated.

best results; these are the Chahine, Landweber, and projection-iterating algorithms. The Chahine algorithm executes most rapidly on the computer, followed by the Landweber and projection-iterating algorithms.

6.3 Adding *a priori* Information

In the previous section I introduced *a priori* information by truncating the matrices to be inverted, so that only the 30 smallest size classes were represented. The assumption implicit was that no particles larger than $x \approx 800$, or $r \approx 80\mu$m, were present. The result was that inversion algorithms converged more rapidly, but offered no additional resolution.

If lower and upper bounds x_{min} and x_{max} on a distribution are known, or can be deduced, these can be used to improve the resolution of an inversion. An example of a distribution with known bounds is the distribution 12dp400 of the previous section. For distributions of particles in nature, such as those in the ocean, an upper bound on a distribution is usually evident.

Upper and lower bounds on x may be appropriate when $d(\theta)$ has clearly not reached a constant within the measured angles. The effect of setting bounds on a distribution is to implicitly assume a value for θ_u, the angle at which $d(\theta)$ is indistinguishable from its asymptote. Rephrasing the Nyquist criterion,

$$\theta_u \leq \frac{\pi(M-1)}{2(x_{max} - x_{min})}. \tag{6.10}$$

Reducing $x_{max} - x_{min}$ increases θ_u. The result is to limit the set of solutions $m(x)$ to those for which $d(\theta)$ approaches its asymptote at a slower rate.

For some distributions where $d(\theta)$ approaches its asymptote slowly,

bounds on x are not appropriate. The function $d(\theta)$ approaches its asymptote slowly for distributions $m(x)$ which contain high frequency content. Distributions with high frequency content contain sharp peaks and discontinuities. Bounds on x may not be appropriate if the function $m(x)$ is a broad distribution, but contains sharp peaks and discontinuities.

For distributions which can be bounded in x, the integration increment Δx implicit in the matrix equation $\mathbf{d} = \mathbf{Gm}$ can be smaller, so that $\mathbf{d} = \mathbf{Gm}$ is a better approximation to its corresponding integral equation. Thus, inversions produce size distributions that fit with less residual error between the measured and predicted scattering.

A size distribution which is both highly peaked and bounded in x is the distribution 12dp400 of the previous section. That distribution is clearly bounded between $x_{min} = 150$ and $x_{max} = 250$. Hence, I present the results of additional inversions, in figure 6.9, which employ these bounds, determining $n(x)$ at 50 sizes.

Bounds on x clearly increase the resolution in inversions of the 40μm particles in the distribution 12dp400. For comparison, I also show the distribution as determined by the supplier, Duke Scientific, via Coulter count measurements. The Coulter count data compares favorably with the inversions. This time, Landweber iteration appears to give the best results, in comparison with the Coulter count.

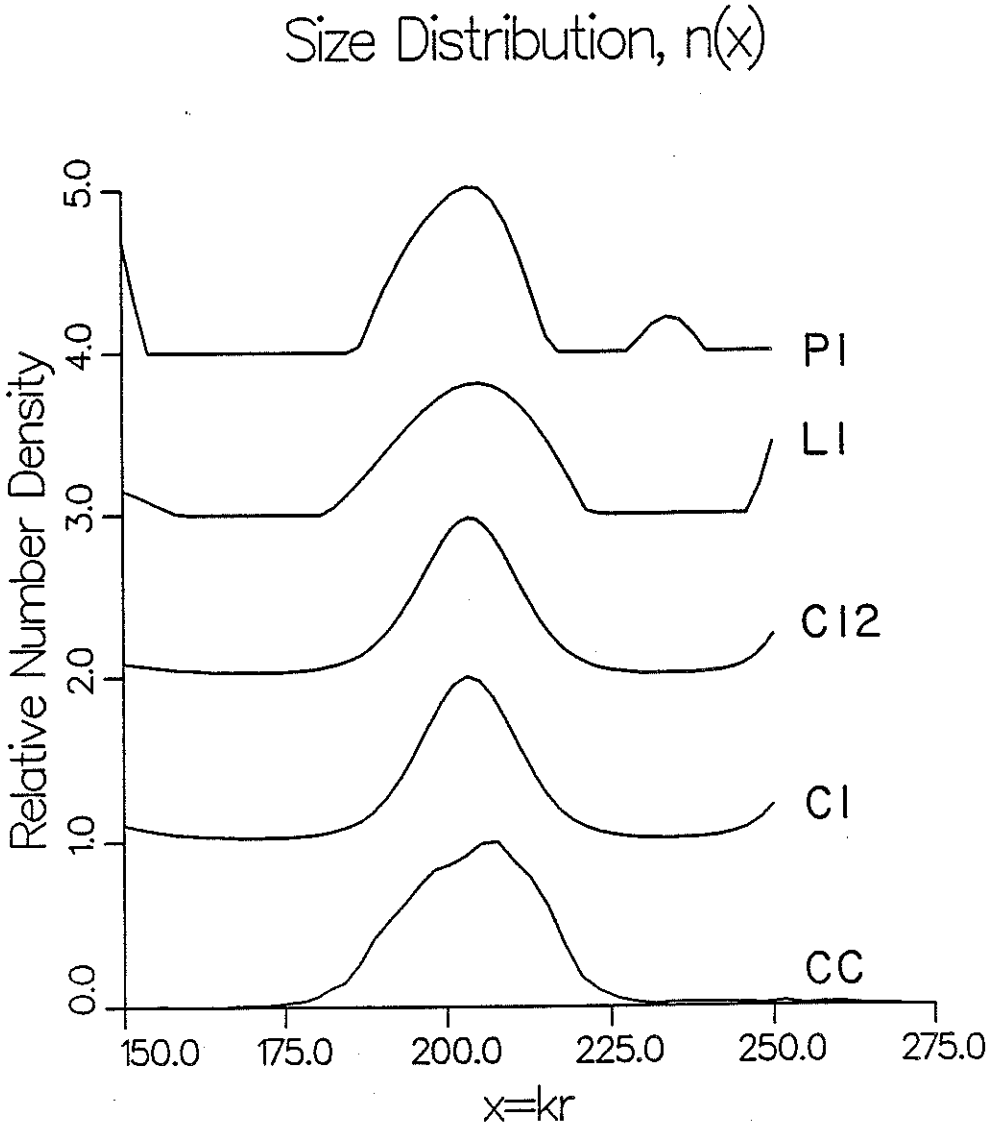

Figure 6.9: Inversions of 12dp400, with $x_{min} = 150$ and $x_{max} = 250$, based on diffraction theory. Also shown are Coulter count measurements. See table 6.5 for legend.

Algorithm (or technique)	Comments	Residual
Projection Iteration (PI)	Worst convergence	0.14
Landweber Iteration (LI)	Closest to Coulter count	0.13
Chahine Iteration (CI2)	80 points resolution in x	0.12
Chahine Iteration (CI)		0.12
Duke Scientic (CC) Coulter Count	Normalized	

Table 6.5: Legend for figure 6.9. Results appear in the same order from top to bottom as in figure 6.5, and labels of curves are shown in parentheses.

6.4 Additional Experimental Results

In this section I present additional results, from measurements of scattering from a multi-modal distribution of particles. These results do not illustrate any additional concepts; they add credence to the previous sections. This distribution was made from three Gaussian distributions, with mean diameters 20.0, 40.3, and 50 μm, or non-dimensional radii $\overline{x} = 100$, 200.15, and 250. For reference, I call this distribution 204050; table 6.6 summarizes its preparation. The relative concentrations of the three species in this distribution were estimated from the approximate concentrations supplied by Duke Scientific.

The observed scattering from the distribution 204050 is shown in figure 6.10; the function $d(\theta)$ is shown in figure 6.11. The random noise for

Name of measurements	204050		
Type of particles	Polystyrene Spheres		
Distribution	Gaussian		
Mean Diameter	20.0μm	40.3μm	50.78μm
Standard Deviation	12%	5.5%	6.9%
Estimated Relative Concentration	6.5	2	1
Source	Duke Scientific		
Catalog Numbers	120	243	244

Table 6.6: Relevant data regarding the distribution named 204050

Name of measurements:	204050	
Number of measurements	97	
Minimum Angle	0.0086	radians
Maximum Angle	0.0446	radians
Number of missing measurements at small angles	22	pixels
Number of CCD scans averaged	400	

Table 6.7: Relevant data regarding the measurement of the distribution 204050.

these measurements is less than 1%. Again, for this data, $d(\theta)$ does not approach a constant asymptote within the range of measurements. This reflects the narrowness of the component size distributions of particles. Relevant data regarding these measurements are contained in table 6.7. Details of the inversions are presented in table 6.8.

Because $d(\theta)$ approaches an asymptote slowly, it is evident that $m(x)$ contains high frequency content; in this case, we know that this is in form

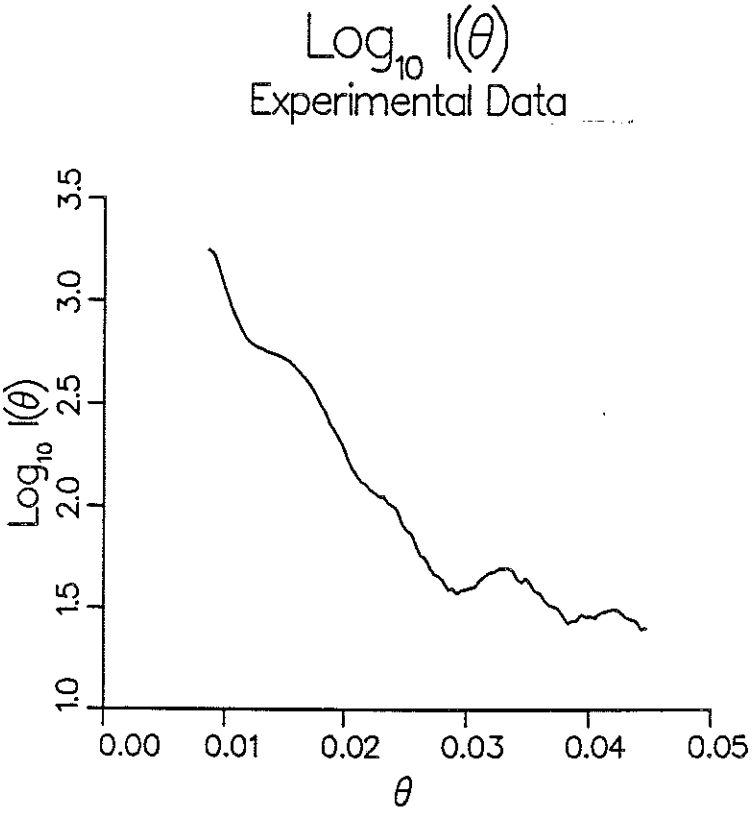

Figure 6.10: Observed Scattering from the tri-modal distribution of polystyrene spheres named 204050.

Name of measurements:	204050
Minimum x inverted for	1.
Maximum x inverted for	300.
Number of sizes inverted for	80
Forward Model for scattering	Diffraction theory

Table 6.8: Information common to all inversions of 204050

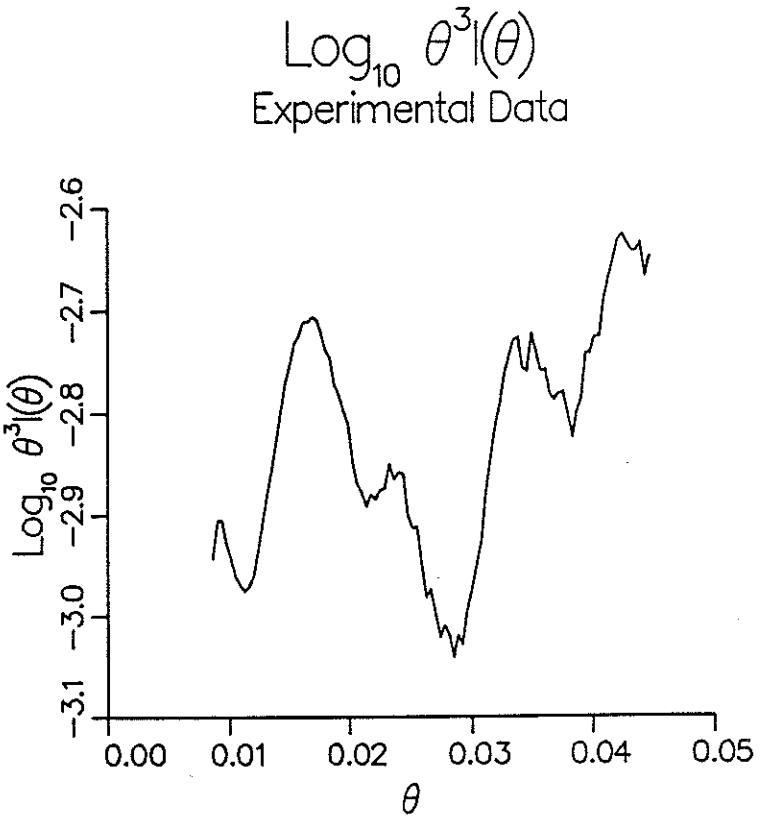

Figure 6.11: $d(\theta)$ for tri-modal distribution of polystyrene spheres named 204050.

Algorithm (or technique)	Comments	Residual
Landweber Iteration (LI)		0.12
Chahine Iteration (CI2)	Same as next, but x =1 to 350.	0.12
Chahine Iteration (CI)		0.13
Chin-Sliepvich Integral Transform (C-S)	Missing Data affects algorithm severely, especially at small x.	
Estimate from Coulter count (CC)		

Table 6.9: Legend for figure 6.12. Labels of curves are shown in parentheses.

of 3 narrow Gaussian peaks. Thus, I found that this data was essentially not invertible without introduction of *a priori* information in the form of bounds on x. By "uninvertible," I mean that convergence of the iterative algorithms was slow, and came to an unacceptably large residual. To speed convergence, I imposed the bounds $x_{min} = 1$ and $x_{max} = 300$, and sought $n(x)$ at 80 sizes.

The results of the inversions of the scattering are shown in figure 6.12. Table 6.9 serves as legend and summary. For comparison, I include an estimate of $n(x)$ based on the size distributions of the 3 component distributions of spheres, supplied by Duke Scientific. I also include an inversion by the Chin-Sliepcevich integral transform.

The results of the inversion algorithms are quite similar, except for the

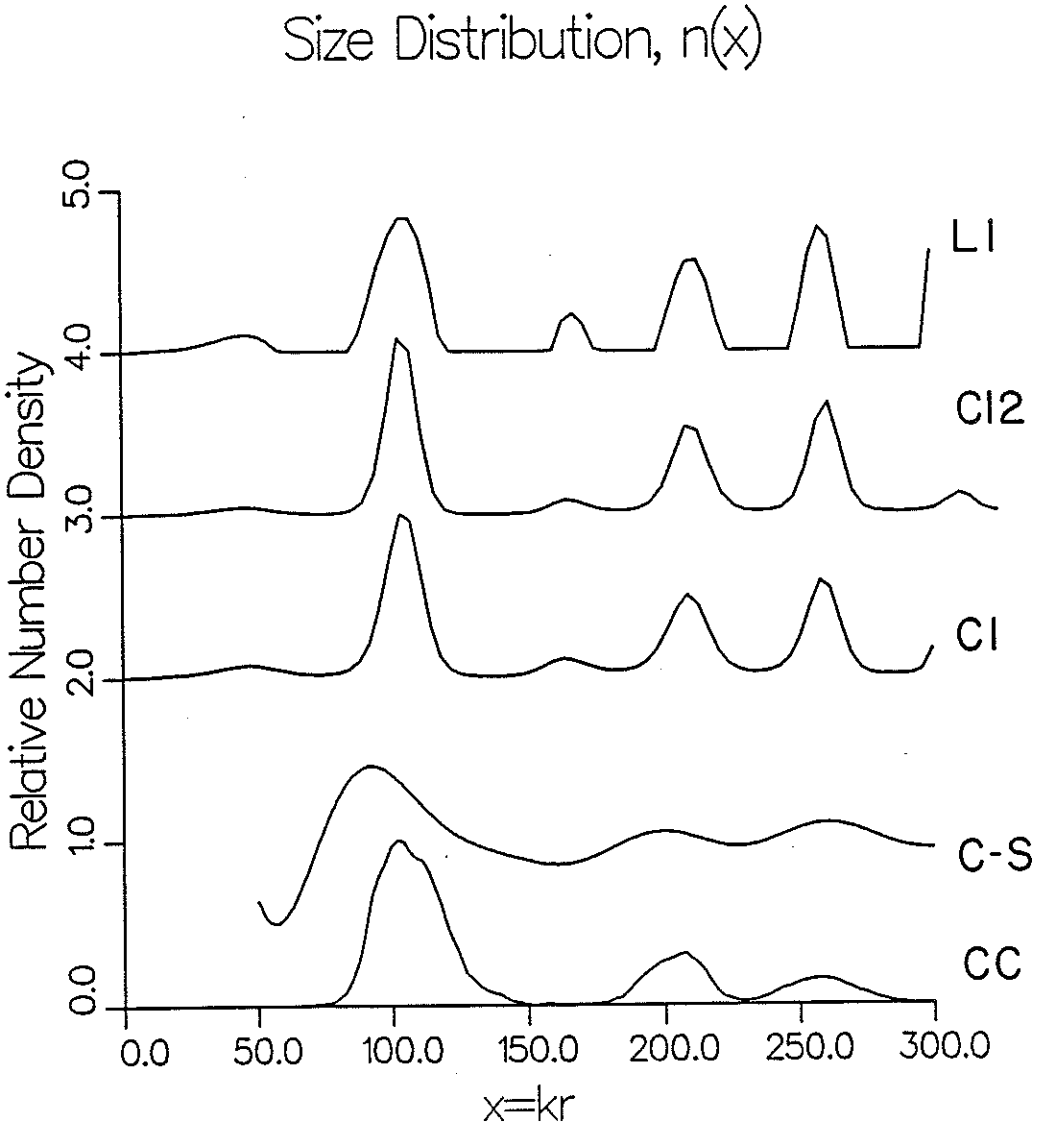

Figure 6.12: Results of inversions, and estimates of size distribution, $n(x)$, for the distribution 204050. Table 6.9 serves as a legend.

Chin-Sliepcevich integral transform, which suffers most from missing data. Comparison to the estimates from Coulter count data is not as good; but this is to be expected, since this is a crude estimate.

6.5 Further Observations: Resolution in Inversions

A notable feature of all the inversion results presented is the lack of high frequency content in $m(x)$. This reflects the limited angular range of the measurements. An algorithm which would discern high frequency content would implicitly make assumptions regarding $d(\theta)$ at angles where it is not measured. In other words, such an algorithm implicitly extrapolates $d(\theta)$. This was discussed in section 3.3. Assuming 100 points of data, with separation 25 μm and an effective focal length of 66.5 mm, the best possible resolution is, according to equation (3.64),

$$\Delta x = \frac{1}{2(\theta_{max} - \theta_{min})} = 13.3 \tag{6.11}$$

This figure compares favorably with the apparent resolution in the inversions.

6.6 Application to Scattering Measurements of Natural Suspensions

We saw that for robust inversion of experimental data, the function $d(\theta)$ should be sampled evenly in the interval $(0, \theta_u]$. Of course, in practice, some data is missing. The reasons for this include insufficient dynamic

range, the short length of the sensing array, and the inability to separate the scattering from the unscattered beam at the smallest angles.

One purpose of this work has been to develop a sampling theory that results in algorithms yielding unique and stable inversions. We have successfully applied the theory to distributions of polystyrene spheres, both mono and poly-dispersed. We have seen that the sampling requirements depend on the size distribution itself.

In an introductory section, we saw that there is a lack of useful data on particle size distributions in the ocean. As a result, it was not possible to apply the theory of sampling directly to known particle size distributions to predict in advance the sampling requirements, in terms of spacing of detectors in θ and in dynamic range. To do so requires experimental measurements of scattering.

The fundamental problems that make complete scattering measurements impossible are insufficient dynamic range, the inability to separate the unscattered beam and the scattering at the smallest angles, and detector arrays that are too short. With this in mind, I present measurements of the scattering data from distributions of particles occurring naturally. These measurements are intended to show the dynamic range necessary for sizing natural distributions. The dynamic range necessary in these measurements is far less than for mono-disperse and multi-modal distributions of polystyrene spheres we have examined. Thus, for these distributions, scattering measurements are available at far smaller angles than for the polystyrene spheres.

We leave inversions of these distributions, which contain both small

($x \approx 1$) and irregularly shaped particles, to later work. This will require determining the appropriate forward scattering model for the particles in question. The appropriate scattering model for both small ($x \approx 1$) and irregular particles are current research topics. de Boer et. al [124] comment that diffraction theory (discussed in chapter 2) does not model the scattering from small ($x \approx 1$) particles accurately. They suggest that Mie theory is more appropriate when particles are small relative to the wavelength.

Furthermore, sizing irregular particles invites, in fact requires, comparison with other particle sizing methods. This requires determining which interpretation of size each method applies to irregular particles. If the particles are to be considered *statistically spherical* in some sense, for example, the sense in which they are statistically spherical must be considered in depth. Consider, for example: what ensemble of spherical particles might be equivalent to a grain of sand, to which a Coulter counter or other particle sizing method might ascribe a certain size? A grain of sand has many characteristic dimensions. The scattering from a sharp point might be represented by many small particles. The radius of a sphere equivalent volume might be represented by a single particle. And, the largest dimension of the particle might be represented by a fractional number of large particles.

I present measured scattering from several natural distributions in figures 6.13, 6.14 and 6.15. These include

- Clay from the HEBBLE site, on the Nova Scotian rise, stirred into solution. This was generously provided by J. F. Lynch and I. N. McCave.

- A sample taken from several feet offshore, in Woods Hole, Mass.
- Vigorously agitated beach sand, in sea water.

The scattering measurements from these distributions was characterized by low noise, typically less than 2 percent of the mean. One can make several observations regarding the scattering from these samples. First, the functions $d(\theta)$ and the scattered intensities from these distributions are smooth; in fact, the sample interval is far smaller than needed. Second, $d(\theta)$ has clearly not approached a constant asymptote. It may be however, that $d(\theta)$ has approached an asymptote; enough particles with sizes close to the wavelength may be present that the $d(\theta)$ is not asymptotic to a constant, as predicted by diffraction theory, but to a constant, plus a function predicted by Rayleigh scattering. Finally, the scattering from these samples requires far less dynamic range to sample than the mono-disperse polystyrene spheres. Because of this, less data is missing at small angles, and there is less noise at large angles. Unlike the mono-disperse spheres, a larger angular range of measurements is easy to obtain for natural suspensions, because less dynamic range is required.

To collect scattering data from oceanic particulates, finally, I draw two conclusions. First, the angular resolution need not be as fine as in the prototype instrumentation (0.0286°). And, the angular range of measurements should be much larger than in the prototype instrument (in which the angular range is approximately 2.8°). The combination of these conclusions may mean that it is possible to measure scattering with no data missing at small angles, resulting in highly stable inversions.

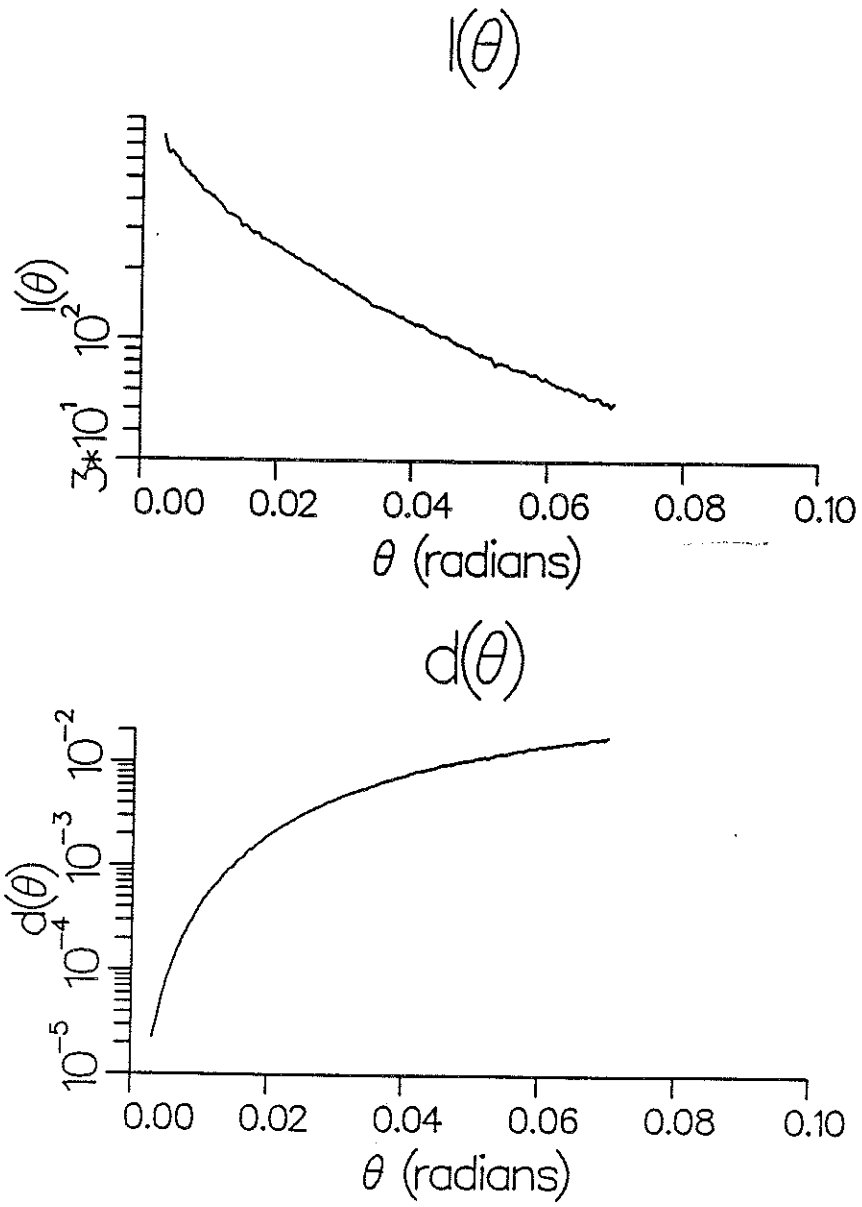

Figure 6.13: Scattering observed from clay from the HEBBLE site, stirred into solution. Top: $I(\theta)$. Bottom: $d(\theta)$. The clay was provided by James F. Lynch and I. N. McCave.

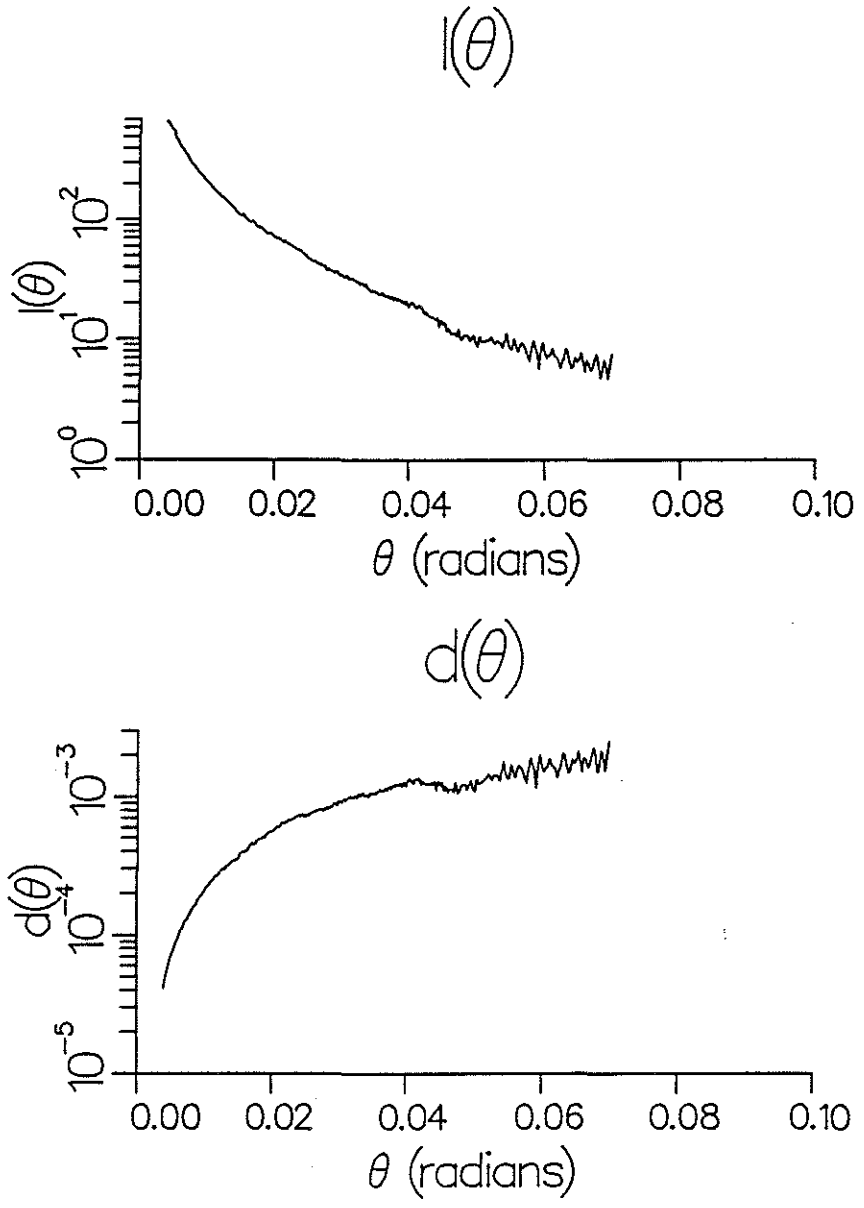

Figure 6.14: Scattering observed from a sample taken several feet offshore, in coastal water of Woods Hole, Mass. Top: $I(\theta)$. Bottom: $d(\theta)$.

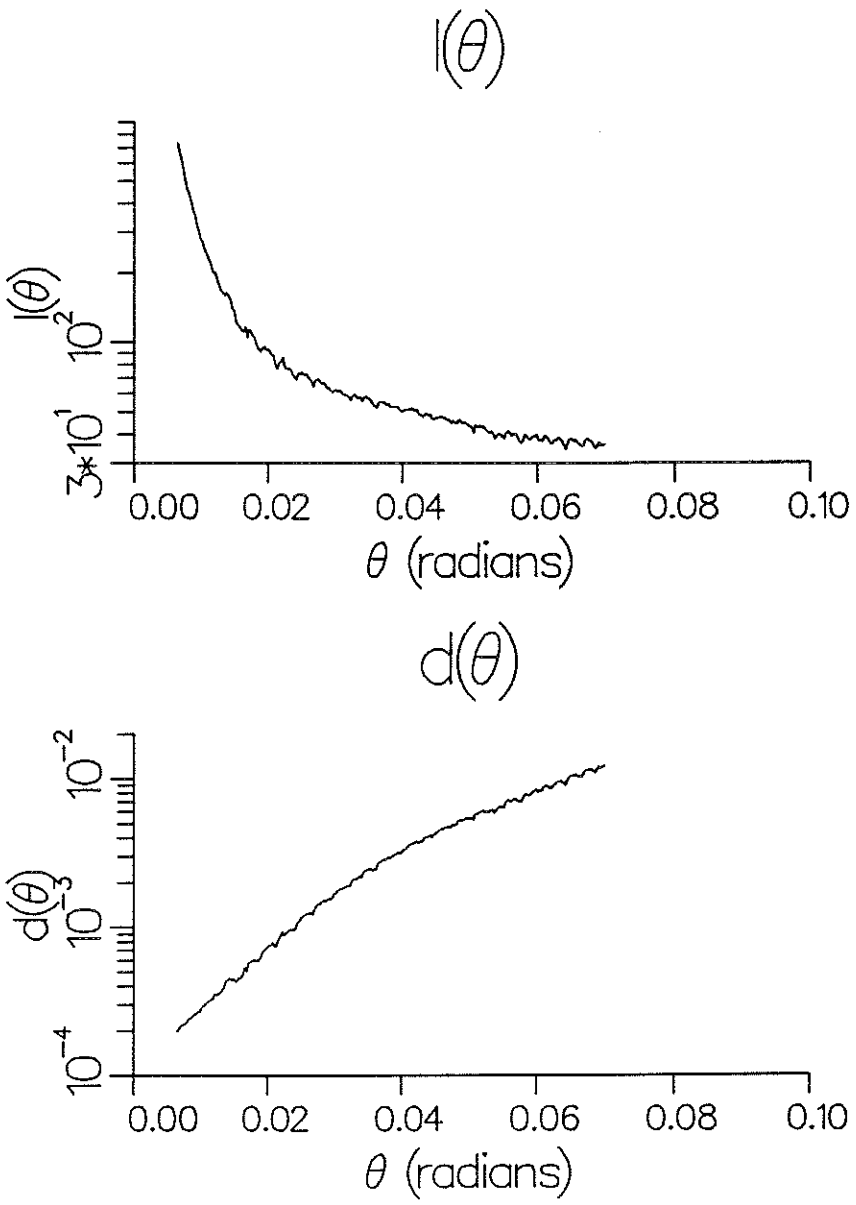

Figure 6.15: Scattering observed from vigorously agitated beach sand, in sea water. Top: $I(\theta)$. Bottom: $d(\theta)$.

Chapter 7

Conclusions

7.1 Contributions

I began this work with the desire to design an optimal sensor for a specific application. Specifically, I sought to design an optical diffraction particle sizing instrument suitable for oceanographic problems. Such an instrument must sample the scattering of a collimated beam, adequately, with sufficiently dense measurements, over an appropriate range of scattering angles. It must have sufficient dynamic range and immunity to noise.

This has led to a study of the sources of non-uniqueness and instability, both mathematical and numerical, of the inverse problem which relates the angular distribution of scattering $I(\theta)$, to the size distribution of suspended particles $n(x)$. This study forms the foundation of this thesis.

Using the asymptotic analysis of the diffraction approximation to scattering from particles, I rephrased the inverse problem, to seek $m(x) = xn(x)$ from $d(\theta) = \theta^3 I(\theta)$. The inversion of $d(\theta)$ to yield $m(x)$ is more stable than of $I(\theta)$ to yield $n(x)$, but is related to it trivially.

Moreover, $d(\theta)$ and $n(x)$ are, except for a constant term, approximately

related by a Fourier-sine transform. Using the Riemann-Lebesgue lemma, I showed that $d(\theta)$ becomes indistinguishable from an asymptote at an angle θ_u justifying sampling $I(\theta)$ only at near forward angles. I derived a sampling criterion, specifying the number of samples between $\theta = 0$ and θ_u required to uniquely represent a given size distribution. If the samples are evenly spaced between $\theta = 0$ and θ_u, but not including $\theta = 0$, the criterion is

$$\Delta\theta \leq \frac{\pi}{2x_u}$$

where x_u is the largest size particle in the distribution. Unevenly spaced samples can also uniquely represent a given size distribution, in principle, as long as the *average* spacing between samples satisfies this criterion. Thus, if the samples are evenly spaced between two angles θ_{min} and θ_{max}, where $\theta_{max} < \theta_u$, the criterion is

$$\Delta\theta \leq \frac{\pi(M-1)}{2x_u}$$

The angle θ_u, at which $d(\theta)$ becomes indistinguishable from an asymptote according to any arbitrary measure, depends on the size distribution. Using the similarity of the relation between $m(x)$ and $d(\theta)$ to a Fourier transform, and the concept of the Heisenberg uncertainty principle, I found a lower bound for θ_u:

$$\theta_u \geq \frac{1}{4x_u}.$$

This is an equality if $m(x)$ is Gaussian.

Though in principle $m(x)$ can be uniquely represented by an uneven sampling of $d(\theta)$, the transformation from $d(\theta)$ to $m(x)$ becomes unstable,

exactly analogous to the instability of determining the Fourier transform of unevenly spaced data. This is important for interpreting inversions of experimental measurements of scattering, in which data is missing at both small and large angles due to insufficient dynamic range, the inability to separate scattering from the direct beam, or detector arrays that are too short. This instability is surmountable too much data is not missing. Inversion algorithms which impose additional constraints on $n(x)$ such as positivity reduce this instability.

Traditionally, the tradeoff between resolution and variance in geophysical inverse problems has been studied using the Backus-Gilbert formalism. This formalism is powerful when its equations are tractable. For diffraction particle sizing, the equations are not tractable. Fortunately, the observation that $d(\theta)$ and $m(x)$ are related by a Fourier sine transform (except for a constant term) reveals that the best resolution attainable, in the presence of noise, is

$$\Delta x = \frac{1}{2(\theta_{max} - \theta_{min})} \tag{7.1}$$

I use the notation

$$\mathbf{d} = \mathbf{Gm}$$

represent the matrix approximation to the integral relation between $d(\theta)$ and $m(x)$. When measurements of $d(\theta)$ are evenly spaced between 0 and θ_u, I showed how to construct G by choosing the sizes x_j judiciously, so that it is approximately orthogonal, in analogy to a matrix which performs a sine transform. With this construction, $\mathbf{d} = \mathbf{Gm}$ represents a discrete sum with a very coarse integration increment, Δx. This increment is small

enough if the approximation

$$\int_0^{x_u} m(x)\,dx = \sum m(i\Delta x)\Delta x$$

is valid.

When measurements of $d(\theta)$ are not evenly spaced, that is, when measurements are missing at the smallest and largest angles, it is no longer possible to construct a matrix that is orthogonal in analogy to a sine transform.

The mathematical instability associated with inverting $d(\theta)$ to yield $m(x)$ shows up in the eigenvalues of the matrix G. The small eigenvalues tend to correspond to small sizes, when G is constructed so that it is approximately orthogonal. This contrasts to the role of eigenvalues in other inverse problems. In other inverse problems, small eigenvalues tend to correspond to high *frequencies* in inversion results. In this problem, however, small eigenvalues correspond not to the frequency content, but to to the *results themselves at small sizes.* As a result, inversion algorithms which manipulate small eigenvalues, such as damped least squares algorithms and singular value decomposition methods, are not suitable for optical diffraction particle sizing. I reviewed other algorithms which do not manipulate small eigenvalues, and which impose other constraints, such as positivity on $n(x)$.

I developed a prototype instrument to measure small angle scattering, to provide experimental data for testing and illustrating the theory, and for making scattering measurements, to empirically determine the sampling requirements for sizing naturally occurring particles. With scattering

data from polystyrene spheres, I compared inversion algorithms; with the missing data typical of experimental measurements, iterative algorithms that impose positivity yield the best results. Other algorithms tend to find erroneous sizes distributions, characterized by negative values.

With the experimental data, I demonstrated the value of *a priori* information in the form of bounds x_{min} and x_{max}. Such bounds are appropriate when θ_{max}, the largest angle sampled, is much smaller than θ_u, the angle at which $d(\theta)$ is indistinguishable from its asymptote.

And finally, I performed measurements of scattering of naturally occurring, irregular particles. These included mud from the HEBBLE site, coastal sea water, from several feet offshore in Woods hole, and beach sand. These revealed that the dynamic range requirements for measuring scattering from polydisperse distributions are less than for the narrow Gaussian size distributed spheres. Further, the resolution of currently available CCD photodiode arrays far exceeds that required to measure scattering from these distributions. With less resolution and measurements over a wider range of angles, it may be possible to measure scattering with no data missing at small angles, resulting in highly stable inversions.

Though I have performed this work, always with an eye toward determining particle size distributions in oceanographic hydrodynamics and light scattering experiments, this work clearly has broader applicability. The work is especially applicable to the sizing of droplet sprays, where particles are spherical, as in combustion and pollution control problems.

7.2 Future Work

I see four areas related to this work for future research. These are

- Development of the forward scattering problem, to model the scattering from irregular particles.
- Inversion algorithm development.
- Improving the prototype scattering instrument, and adapting it for use in oceanographic work.

In chapter 2, I discussed several forward models. These models are sufficient for much of this work, in that they accurately predict the frequency content and asymptotic tendencies of the scattered intensity, $I(\theta)$. To invert the scattering from distributions of irregular particles, however, better models of scattering may be required. Chylek [79,125] has performed much of the research on light scattering from irregular particles.

To obtain better models of forward scattering from irregular particles, one might envision an experiment, in which the scattering from an ensemble of irregular particles would be measured. Statistics regarding the shapes of these particles would then be obtained from scanning electron micrography. These in turn would be used to compute scattering, using Fraunhoffer diffraction theory. The scattering is then the magnitude of the two dimensional Fourier transform of an ensemble of statistically typical particles. The results could be compared to the work Jaggard and Kim reported [134]. These investigators use fractals to describe the diffraction of light by irregular structures.

In this work, I compared a number of commonly used inversion algorithms; I did not develop new algorithms. One can, however, envision new algorithms, which would not only invert measurements, but extrapolate them where data is missing, at the same time constraining $n(x)$ to be positive, and $I(\theta)$ to be positive, and bounded, or band-limited. Such an algorithm would be modelled after algorithms applied to extrapolation of band-limited signals and super-resolution. [99,126] These are iterative algorithms. For this problem, an algorithm would work as follows:

1. One would start with experimental data, $\mathbf{d}$ which would be missing data at the smallest and largest angles. One would use this to form a new array of data, $\mathbf{d}'$, in which the missing data would be filled in with guesses.

2. The data would be inverted, using a conventional algorithm, yielding an estimate $\hat{\mathbf{m}}$.

3. $\hat{\mathbf{m}}$ would be "passed" through a constraint operator, which would set negative number densities to zero, and perhaps constrain the size distribution to lie between two sizes.

4. A new intensity $\mathbf{d}'$ would be calculated from the size distribution.

5. $\mathbf{d}'$ would be passed through a constraint operator, setting negative intensities to zero, and insuring that intensities are bounded. Most importantly, where experimental data is not missing, $\mathbf{d}'$ would be reset to equal the experimental data, $\mathbf{d}$.

6. The iteration returns to step 2 until convergence is reached.

Linear programming may also serve as an algorithm for inverting small angle scattering; this was discussed in chapter 4.

I have discussed improvements to the prototype instrument for measuring small angle forward scattering in section 5.2. These improvements are directed primarily toward improving the dynamic range of the measurements, to obtain data at the smallest angles, where intensity is highest, and the largest angles, where it is least. For oceanographic experiments, the instrument will have to be packaged in a pressure case. For experiments in which forced stirring of the sample volume is not desirable, a detector which averages scattered intensity spatially, rather than temporally will be desirable, to eliminate speckle noise. Such an experiment might measure particle fall velocity, for example.

Finally, much of the theory in this thesis can be applied to the related problem of inverting spectral, rather than angular measurements of scattering. This is touched upon in appendix B.

Appendix A

Derivation of the Chin-Sliepcevich Inverse

Here we derive equation (3.3), the Chin-Sliepcevich inverse, which is

$$n(x) = \frac{-2\pi k^3}{x^2} \int_0^\infty J_1(x\theta) Y_1(x\theta) x\theta \, \frac{d}{d\theta} \left[\theta^3 I(\theta)\right] \, d\theta \tag{A.1}$$

and is the inverse of

$$I(\theta) = \int_0^\infty J_1^2(x\theta) \, x^2 n(x) \, dx \tag{A.2}$$

The Chin-Sliepcevich inverse is actually not essential to our work; results in this thesis do not rely on it. Nonetheless, it has served as the basis of nearly all prior studies of uniqueness, instability and sampling for the optical diffraction sizing problem [127,11,10,12,13]. This derivation closely follows Fymat's [10].

The derivation begins with the Bateman-Titchmarsh-Fox transform:

$$\phi(t) = 2\pi \int_0^\infty \frac{d}{dt}[tJ_\nu^2(rt)] \, rf(r) \, dr, \tag{A.3}$$

and

$$f(s) = -\int_0^\infty J_\nu(st) Y_\nu(st) \, t\phi(t) \, dt, \tag{A.4}$$

which applies when $\nu \geq 0$ and $rf(r)$ and $t(\phi(t)$ are integrable over $(0, \infty)$. The transform is actually a generalization of the Fourier-Sine transform, which it reduces to for $\nu = \pm 1/2$.

Equation (A.2) can be written

$$F(\theta) = (2\pi k^3 \theta^3) I(\theta) = 2\pi \int_0^\infty \theta J_1^2(x\theta)\, x f(x)\, dx \qquad \text{(A.5)}$$

with the introduction of $F(\theta)$ and $f(x) = xn(x)$. Differentiating under the integral sign, this becomes

$$\frac{d}{d\theta} F(\theta) = 2\pi \int_0^\infty \frac{d}{d\theta} [\theta J_1^2(x\theta)]\, x f(x)\, dx \qquad \text{(A.6)}$$

This is in the form of equation (A.3), with $\nu = 1$ and $\phi(t) = d/d\theta F(\theta)$. Thus the transform becomes

$$f(x) = -\int_0^\infty J_1(x\theta) Y_1(x\theta)\, \frac{d}{d\theta} F(\theta)\, d\theta \qquad \text{(A.7)}$$

which immediately yields equation (3.3).

Appendix B

Inversion of Spectral Measurements

Some investigators are interested in inverting *spectral* measurements, rather than angular measurements [128,129,130,87,131,132,133]. Much of the theory in this thesis is applicable. We saw that for $x_u\theta > 4$,

$$I(\theta) \approx \frac{1}{\pi\theta^3 k^3}\left[\int_0^\infty xn(x)\,dx - \int_0^\infty \sin 2x\theta\, xn(x)\,dx\right] \tag{B.1}$$

Revising notation, this becomes

$$I(k) \approx \frac{1}{\pi\theta^3 k}\left[\int_0^\infty rn(r)\,dr - \int_0^\infty \sin 2kr\theta\, rn(r)\,dr\right] \tag{B.2}$$

Where the intensity is now a function of wave-number and θ is just a parameter. Setting $z = r\theta$, $M(z) = zn(z)$, and $D(k) = kI(k)$,

$$D(k) = \frac{1}{\pi\theta^5}\left[\int_0^\infty zn(z)\,dz - \int_0^\infty \sin 2zk\, M\,dz\right] \tag{B.3}$$

In this case, the (dimensional) variables z and k are conjugate exactly the way the (non-dimensional) variables θ and x are. In fact, λ, θ, and x are all mutually conjugate. Now, θ is just a parameter, although for validity, one must have $\theta \ll 1$.

Having recognized that z and k are conjugate, sampling criteria, resolution, etc. can be derived analogously, with some caveats. One caveat is

that $D(k)$ can be measured only over a narrow range of large k values, resulting in unstable inversions. Furthermore, the effect of resonances in the scattering becomes a greater problem, since, to first order, equation (B.1) differs from real scattering by a constant factor. This factor is wave-length dependent.

Appendix C

Frequently Used Variables

$n(x)$	Particle number density
x	Particle Radius, non-dimensional; $x = kr$
r	Particle Radius
k	Wave number, $2\pi/\lambda$
M	Number of samples in θ
N	Number of samples in x
$J_1(y)$	Bessel function of first kind, first order
$Y_1(y)$	Bessel function of second kind, first order
θ	Scattering angle
$I(\theta)$	Scattered Intensity
$I(\theta, x)$	Scattering from particle of radius x at angle θ
$d(\theta)$	$\theta^3 I(\theta)$
$m(x)$	$xn(x)$
x_u	Largest size in a distribution
θ_{valid}	Largest θ at which diffraction approx. is valid
θ_u	θ at which $d(\theta)$ is undistinguishable from asymptote
θ_{min}	Minimum sampling angle
θ_{max}	Maximum sampling angle
θ_i	Discrete angles at which sampling occurrs
$G(\theta, x)$	$\theta^3 I(\theta, x)/x$

Table C.1: Frequently used variables and functions

References

[1] H. R. Gordon, R. C. Smith, and J. R. V. Zanefeld. Introduction to ocean optics. In Marvin A. Blizzard, editor, *OCEAN OPTICS VII*, pages 2–41, SPIE, 1984.

[2] I. N. McCave. Vertical flux of particles in the ocean. *Deep Sea Res.*, 22:491–502, 1975.

[3] C. E. Lambert, C. Jehanno, N. Silverberg, Brun-Cottoan, and R. Chesselet. Log-normal distributions of suspended particles in the open ocean. *J. Mar. Res.*, 39:77–98, 1981.

[4] Azzopardi. The analysis of Malvern ST1800 output by different models and a comparative test of a near monodisperse distribution. 1980.

[5] L. Brecevic and J. Garside. On the measurement of crystal size distributions in the micrometer size range. *Chem. Eng. Sci.*, 36:867–869, 1981.

[6] D. C. Hammond. Accuracy verification of a Malvern ST-1800 analyzer.

[7] I. N. McCave. Evaluation of a Laser-Diffraction-Size analyzer for use with natural sediments. *J. Sedimentary Petrology*, 56:561–564, 1986.

[8] J. H. Chin, C. M. Sliepcevich, and M. Tribus. Particle size distributions from angular variation of forward-scattered light at very small angles. *J. Phys. Chem.*, 59:841–844, 1955.

[9] J. H. Chin, C. M. Sliepcevich, and M. Tribus. Determination of particle size distributions by means of measurements of angular variation of intensity of forward-scattered light at very small angles. *J. Phys. Chem.*, 59:845–848, 1955.

[10] A. L. Fymat and K. D. Mease. Reconstructing the size distribution of spherical particles from angular forward scattering data. In A. L. Fymat and V. E. Zuev, editors, *Remote Sensing of the Atmosphere: Inversion Methods and Applications*, pages 195–231, Elsevier, Amsterdam, 1978.

[11] J. H. Koo and E. D. Hirleman. Comparative study of laser diffraction analysis using integral transform technqiues: factors affecting the reconstruction of droplet size distributions. In *1986 Joint Meeting, Canadian and Western State Sections*, 1986.

[12] K. S. Shifrin. The essential range of scattering angles in measuring particle-size distribution by the small-angle method. *Izv. Atmospheric and Oceanic Phys.*, 2:559–561, 1966.

[13] K. S. Shifrin and I. B. Kolmakov. Effect of limitation of the range of measurement of the indicatrix on the accuracy of the small-angle method. *Izv. Atmospheric and Oceanic Phys.*, 2:514–518, 1966.

[14] K. S. Shifrin, O. V. Kopelvich, V. I. Burenkov, and Y. L. Mashtakov. *Using the Light Scattering Function for Studying Suspended Matter in the Sea.* Technical Report, NASA, 1973.

[15] R. N. Sokolov, F. A. Kudryavitskiy, and G. D. Petrov. Submarine laser instrument for measuring the size spectra of particles suspended in sea water. *Izv. Atmospheric and Oceanic Phys.*, 7:673–675, 1971.

[16] J. N. Hunt. On the turbulent transport of a heterogeneous sediment. *Quart. J. Mech. Appl. Math.*, 22:235–246, 1969.

[17] J. N. Hunt. The turbulent transport of suspended sediment in open channels. *Proc. R. Soc. London A*, 224:322–335, 1954.

[18] P. C. Benedict *et. al.* Sediment transportation mechanics: suspension of sediment. *J. Hydraulics Div. Proc. ASCE*, 89,(HY5):45–76, 1963.

[19] W. D. Grant and S. Glenn. *Continental Shelf Bottom Boundary Layer Model: Volume I: Theoretical Model Development.* Technical Report, American Gas Association, 1983.

[20] J.D. Smith. *The Sea.* Volume 6, J. Wiley and Sons, New York, 1977.

[21] C. E. Adams Jr. and G. L. Weatherly. Some effects of suspended sediment stratification on an oceanic bottom boundary layer. *J. Geophys. Res.*, 86:4161–4172, 1981.

[22] W. H. Wells. Loss of resolution as a result of multiple small-angle scattering. *J. Opt. Soc. Am.*, 53:214, 1963.

[23] V. A. Del Grosso. Optical transfer function measurements in the Sargasso sea. In *Ocean Optics V*, pages 74–101, SPIE, 1975.

[24] C. E. Junge. *Air Chemistry and Radioactivity.* Academic Press, New York, 1963.

[25] K. L. Carder, G. F. Beardsley, and H. Pak. Particle size distributions in the eastern Equatorial Pacific. *J. Geophys. Res.*, 76:5070–5077, 1971.

[26] E. T. Baker and G. A. Cannon. Particle transport processes in a small marine bay. *J. Geophys. Res.*, 88:9661–9669, 1983.

[27] R. W. Sheldon. A continuous size spectrum for particulate matter in the sea. *J. Fish. Res. Bd. Canada*, 24:909–915, 1967.

[28] H. Bader. The hyperbolic distribution of particle sizes. *J. Geophys. Res.*, 75:2822–2830, 1970.

[29] R. W. Sheldon, A. Prakash, and W. H. Sutcliffe, jr. The size distribution of particles in the ocean. *Limnol. Oceanogr.*, 17:327–340, 1972.

[30] O. B. Brown and H. R. Gordon. Two component Mie scattering models of Sargasso sea particles. *Appl. Opt.*, 12:2461–2465, 1973.

[31] D. Lal and A. Lerman. Size spectra of biogenic particles in ocean water and sediments. *J. Geophys. Res.*, 80:423–430, 1975.

[32] J. K. Bishop, D. R. Ketten, and J. M. Edmond. The chemistry, biology and vertical flux of particulate matter from the upper 400 m of the Cape Basin in the southeast Atlantic Ocean. *Deep Sea Res.*, 25:1121–1161, 1978.

[33] H. Pak, J. R. V. Zanefeld, and G. F. Beardsley. Mie scattering by suspended clay particles. *J. Geophys. Res.*, 76:5065–5069, 1971.

[34] R. W. Spinrad, J. R. V. Zanefeld, and J. C. Kitchen. A study of the optical characteristics of the suspended particles in the benthic nepheloid layer of the scotian rise. *J. Geophys. Res.*, 88:7641–7645, 1983.

[35] R. Tsuda and K. Nakata. Particle size distribution and light scattering in Akita bay. *La Mer*, 20:1–8, 1982.

[36] O. B. Brown and H. R. Gordon. Size-refractive index distribution of clear coastal water particulates from light scattering. *Appl. Opt.*, 13:2874–2881, 1974.

[37] H. R. Gordon and O. B. Brown. A theoretical model of light scattering by Sargasso sea particulates. *Limnology and Oceanography*, 17:826–832, 1972.

[38] S. Sugihara and R. Tsuda. Light scattering and size distribution of particles in the surface waters of the north Pacific ocean. *J. Oceanogr. Soc. Japan*, 35:82–90, 1979.

[39] J. R. V. Zanefeld, D. M. Roach, and H. Pak. The determination of the index of refraction distribution of oceanic particulates. *J. Geophys. Res.*, 79:4091–4095, 1974.

[40] D. S. Ensor and M. J. Pilat. The effect of particle size distribution on light transmittance measurement. *Amer. Indust. Hygiene Assoc.*

Jour., 32:287–292, 1971.

[41] E. D. Hirleman. *Non-Intrusive Laser Particle Diagnostics.* Technical Report, AIAA, 1983.

[42] D. J. P. Swift, J. R. Schubel, and R. W. Sheldon. Size analysis of fine-grained suspended sediments: a review. *J. Sedimentry Petrology*, 42:122–134, 1972.

[43] B. J. Thompson, J. H. Ward, and W. R. Zinky. Application of hologram techniques for particle size analysis. *Appl. Opt.*, 6:519–526, 1967.

[44] K. L. Carder, R. G. Steward, and P. R. Betzer. In situ holographic measurements of the sizes and settling rates of oceanic particulates. *J. Geophys. Res.*, 87:5681–5685, 1982.

[45] J. R. V. Zanefeld, R. W. Spinrad, and R. Bartz. An optical settling tube for the determination of particle size distributions. *Marine Geology*, 49:357–376, 1982.

[46] W. M. Farmer. Measurement of particle size, number density, and velocity using a laser interferometer. *Appl. Opt.*, 11:2603–2612, 1972.

[47] P. C. Ariesohn, S. A. Self, and R. H. Eustis. Two-wavelength laser transmissometer for measurements of the mean size and concentration of coal ash droplets in combustion flows. *Appl. Opt.*, 19:3775–3781, 1980.

[48] W. D. Bachalo. Method for measuring the size and velocity of spheres by dual-beam light-scatter interferometry. *Appl. Opt.*, 19:363–370, 1980.

[49] Gaskill. *Linear Systems, Fourier Transforms, and Optics.* J. Wiley and Sons, New York, 1978.

[50] J. W. Goodman. *Introduction to Fourier Optics.* McGraw-Hill, New York, 1968.

[51] W. L. Anderson and R. E. Beissner. Counting and classifying small objects by far-field light scattering. *Appl. Opt.*, 10:1503–1508, 1971.

[52] W. L. Anderson and S. Y. Shen. Statistical inference of particle populations from scattered light. *J. Opt. Soc. Am.*, 69:1684–1690, 1979.

[53] L. G. Dodge. Change of calibration of diffraction-based particle sizers in dense sprays. *Optical Engineering*, 23:626–630, 1984.

[54] E. D. Hirleman. *On-Line Calibration for Laser Diffraction Droplet Sizing Instruments.* Technical Report, American Society of Mechanical Engineers, 1983.

[55] E. D. Hirleman, V. Oechsle, and N. A. Chigier. Response characteristics of laser diffraction particle size analyzers: optical sample volume extent and lens effects. *Opt. Eng.*, 23:610–619, 1984.

[56] D. Holve and S. A. Self. Optical particle sizing for *in situ* measurements, part 1. *Appl. Opt.*, 18:1632–1645, 1979.

[57] D. Holve and S. A. Self. Oicle sizing for *in situ* measurements, part 2. *Appl. Opt.*, 52, 1979.

[58] H. Stark, D. Lee, and B. Vounting quasi-circular particles by an optical-digital metho*pt.*, 15:2246–2249, 1976.

[59] H. Stark and G. Shao. Dderations in counting particles by size with an optical-digi; method. *Appl. Opt.*, 16:1670–1674, 1977.

[60] J. Swithenbank, J. M. BeTaylor, D. Abbot, and G. C. McCreath. *A Laser Diagh̄nique for the Measurement of Droplet and Particle Sizeion.* Technical Report 76-79, AIAA, 1976.

[61] B. M. Tishkoff. Spray chamn: practices and requirements. *Opt. Eng.*, 23:557–560, 198

[62] C. G. Bohren and D. R. Hu*sorption and Scattering of Light by Small Particles.* John Wons, New York, 1983.

[63] L. P. Bayvel and A. R. J*tromagnetic Scattering and its Applications.* Applied Scieners, London, 1981.

[64] H. C. van de Hulst. *Light by Small Particles.* Dover Publications, Inc., New York, 1

[65] M. Kerker. *The Scattering d Other Electromagnetic Radiation.* Academic Press, Lon

[66] G. Grehan and G. Gouesbet. Mie theory calculations: new progress, with emphasis on particle sizing. *Appl. Opt.*, 18:3489–3493, 1979.

[67] W. J. Lentz. Generating Bessel functions in Mie scattering using continued fractions. *Appl. Opt.*, 15:668–671, 1976.

[68] J. B. Riley and Y. C. Agrawal. *A VAX/VMS FORTRAN Routine For Mie Scattering.* Technical Report, WHOI, 1987.

[69] W. J. Wiscombe. Improved Mie scattering algorithms. *Appl. Opt.*, 19:668–671, 1980.

[70] B. Verner. *J. Opt. Soc. Am.*, 66:1424–1425, 1976.

[71] H. A. Haus. *Waves and Fields in Opto-Electronics.* Prentice-Hall, New York, 1982.

[72] A. R. Jones. Error contour charts relevant to particle sizing by forward-scattered lobe methods. *J. Phys. D*, 10:L163–L165, 1977.

[73] S. Boron and B. Waldie. Particle sizing by forward lobe scattered intensity-ratio technique: errors introduced by applying diffraction theory in Mie regime. *Appl. Opt.*, 17:1644–1648, 1978.

[74] A. L. Fymat. Analytical inversions in remote sensing of particle size distributions. 2: angular and spectral scattering in diffraction approximations. *Appl. Opt.*, 17:1677–1678, 1978.

[75] A. L. Fymat. Remote sensing of environmental particulate pollutants: optical methods for determinations of size distribution and

complex refractive index. In D. J. Collins, M. S. Plesset, and M. M. Saffren, editors, *Proceedings of the International Colloquium on Drops and Bubbles*, page 572, U.S. Government, 1974.

[76] A. L. Fymat and K. D. Mease. Mie forward scattering: improved seMiempirical approximation with application to particle size distribution inversion. *Appl. Opt.*, 20:194–198, 1981.

[77] G. O. Levine, S. Olaofe. Scattering of electromagnetic waves by two equal spherical particles. *J. Coll. Interface Sci.*, 27:442–457, 1968.

[78] S. Asano and G. Yamamoto. Light scattering by a spheroidal particle. *Appl. Opt.*, 14:29–49, 1975.

[79] P. Chylek, G. W. Grams, and R. G. Pinnick. Light scattering by irregular randomly oriented particles. *Science*, 193:480–482, 1976.

[80] C. Acquista. Validity of modifying Mie theory to describe scattering by nonspherical particles. *Appl. Opt.*, 17:3851–3852, 1978.

[81] R. J. Gibbs. Light scattering from particles of different shapes. *J. Geophys. Res.*, 83:501–502, 1978.

[82] R. G. Pinnick, D. E. Carroll, and D. J. Hofmann. Polarized light scattered from monodisperse randomly oriented particles: measurements. *Appl. Opt.*, 15:384–393, 1976.

[83] D. W. Shuerman, R. T. Wang, and R. W. Gufstafson, Schaefer. Systematic studies of light scattering. 1: particle shape. *Appl. Opt.*, 23:4039–4050, 1981.

[84] J. W. Goodman. Statistical properties of laser speckle patterns. In J. C. Dainty, editor, *Laser Speckle and Related Phenomena*, pages 11–75, Springer-Verlag, Berlin, 1984.

[85] D. D. Jackson. Interpretation of inaccurate, insufficient and inconsistent data. *Geophys. J. R. Astr. Soc.*, 28:97–109, 1972.

[86] J. V. Dave. Determination of size distribution of spherical polydispersions using scattered radiation data. *Appl. Opt.*, 10:2035–2044, 1971.

[87] E. R. Westwater and A. Cohen. Application of Backus-Gilbert inversion technique to determination of aerosol size distributions from optical scattering measuremtents. *Appl. Opt.*, 12:1340–1348, 1973.

[88] M. Z. Hansen. Atmospheric particulate analysis using angular light scattering. *Appl. Opt.*, 19:3441–3448, 1980.

[89] A. N. Tikononv and V. Y. Arsenin. *Solutions of Ill-Posed Problems.* John Wiley and Sons, New York, 1977.

[90] A. W. Naylor and G. R. Sell. *Linear Operator Theory in Engineering and Science.* Volume 40 of *Applied Mathematical Sciences Series*, Springer–Verlag, New York, 1982.

[91] C. M. Bender and S. A. Orzag. *Advanced Mathematical Methods for Scientists and Engineers.* McGraw–Hill, New York, 1978.

[92] A. J. Jerri. The Shannon sampling theorem – its various extensions and applications: a tutorial review. *Proc. IEEE*, 65:1565–1596, 1977.

[93] J. L. Yen. On nonuniform sampling of bandwidth-limited signals. *IEEE Trans. Circuit Theory*, IT-6:251–257, 1956.

[94] J. V. Dave. Effect of coarseness of the integration increment on the calculation of the radiation scattered by polydispersed aerosols. *Appl. Opt.*, 8:1161–1167, 1969.

[95] D. Slepian and H. O. Pollak. Prolate spheroidal wave functions, Fourier analysis and uncertainty — I. *Bell Sys. Tech. J.*, 40:43–63, 1961.

[96] G. A. Viano. On the extrapolation of optical image data. *J. Math. Phys.*, 17:1160–1165, 1976.

[97] M. S. Sabri and W. Steenhaart. An approach to band–limited extrapolation: the extrapolation matrix. *IEEE Trans. on Circ. and Sys.*, CAS-25:74–78, 1978.

[98] J. Maeda and K. Murata. Restoration of band-limited images by an iterative regularized psuedoinverse method. *J. Opt. Soc. Am. A*, 1:28–34, 1984.

[99] A. Papoulis. A new algorithm in spectral analysis and band-limited extrapolation. *IEEE Trans. Circuits and Systems*, CAS-22:735–742, 1975.

[100] D. C. Youla. Generalized image restoration by the method of alternating orthogonal projections. *IEEE Trans. Circuits and Systems*, CAS-25:694–702, 1978.

[101] G. M. Jenkins and D. G. Watts. *Spectral Analysis and its Applications.* Holden-Day, Oakland, California, 1968.

[102] D. G. Luenberger. *Optimization by Vector Space Methods.* John Wiley and Sons, New York, 1969.

[103] F. Hagin. On the construction of well-conditioned systems for Fredholm I problems by mesh adapting. *J. Comp. Phys.*, 36:154–169, 1980.

[104] A. V. Oppenheim and R. W. Schafer. *Digital Signal Processing.* Prentice Hall, Inc., Englewood Cliffs, New Jersey, 1975.

[105] A. V. Oppenheim, A. S. Wilskey, and I. T. Young. *Signals and Systems.* Prentice Hall, Inc., Englewood Cliffs, New Jersey, 1983.

[106] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions.* Dover Publications, New York, ninth edition, 1970.

[107] C. Lanczos. *Linear Differential Operators.* D. Van Nostrand Co., London, 1961.

[108] R. A. Wiggins. The general linear inverse problem: implication of surface waves and free oscillations for earth structure. *Rev. Geophys. and Space Phys.*, 10:251–283, 1972.

[109] J. N. Franklin. Well-posed stochastic extensions of ill-posed linear problems. *J. Math. Anal. Appl.*, 31:682–716, 1970.

[110] K. Aki and P. G. Richards. *Quantitative Seismology, Theory and Methods.* Volume II, W. H. Freeman and Co., San Francisco, 1980.

[111] D. W. Oldenberg. Calculation of Fourier transforms by the Backus-Gilbert method. *Geophys. J. R. astr. Soc.*, 44:413–431, 1976.

[112] R. W. Schafer, R. M. Mersereau, and M. A. Richards. Constrained iterative restoration algorithms. *Proc. IEEE*, 69:432–450, 1981.

[113] S. Twomey. *Introduction to the Mathematics of Inversion in Remote Sensing and Indirect Measurements.* Elsevier, Amsterdam, 1977.

[114] T. S. Huang, D. A. Barker, and S. P. Berger. Iterative image restoration. *Appl. Opt.*, 14:1165–1168, 1975.

[115] C. L. Lawson and D. J. Hanson. *Solving Least Squares Problems.* Prentice-Hall, Englewood Cliffs, New Jersey, 1974.

[116] W. Menke. *Geophysical Data Analysis: Discrete Inverse Theory.* Academic Press, New York, 1984.

[117] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes: The Art of Scientific Computing.* Cambridge University Press, Cambridge, 1986.

[118] P. C. Sabatier. Positivity constraints in linear inverse problems — I., general theory. *Geophys. J. R. astr. Soc.*, 48:415–441, 1977.

[119] *Image Sensing Products.* EG&G Reticon, Sunnyvale, Calif.

[120] D. H. Sheingold. *Analog - Digital Conversion Notes.* Analog Devices, Inc., Norwood, Massachusetts, 1980.

[121] R. W. Simpson. Noise in large-aperture self-scanned diode arrays. *Rev. Sci. Instrum.*, 50:730–732, 1979.

[122] Y. Talmi and R. W. Simpson. Self-scanned photodiode array: a multichannel spectrometric detector. *Appl. Opt.*, 19:1401, 1980.

[123] Y. C. Agrawal and J. B. Riley. Optical particle sizing for hydrodynamics based on near forward scattering. *OCEAN OPTICS VII*, Marvin A. Blizzard, editor, Proc. SPIE 489, pages 68–76, 1984.

[124] G. B. J. de Boer, C. de Weerd, D. Thoenes, and W. J. Goossens. Laser diffraction spectromentry: Fraunhoffer diffraction versus Mie scattering. *Part. Charact.*, 8:14–19, 1987.

[125] P. Chylek and V. Ramaswamy. Lower and upper bounds on extinction cross sections of arbitrarily shaped strongly absorbing or strongly reflecting nonspherical pariticles. *Appl. Opt.*, 21:4339–4344, 1982.

[126] R. J. Marks II. Gerchberg's extrapolation in two dimensions. *Appl. Opt.*, 20:1815–1820, 1981.

[127] J. N. Desai and D. B. Vaidya. Particle size analysis by forward scattering. *J. Coll. Interface Sci.*, 51:527–531, 1975.

[128] M. D. King, D. M. Byrne, B. M. Herman, and J. A. Reagan. Aerosol size distributions obtained by inversion of spectral optical depth measurements. *J. Atmos. Sci.*, 35:2153–2167, 1978.

[129] A. L. Fymat. Analytical inversions in remote sensing of particle size distributions. 1: multispectral extinctions in the anomalous diffraction approximation. *Appl. Opt.*, 17:1675–1676, 1978.

[130] A. L. Fymat. Analytical inversions in remote sensing of particle size distributions. 3: angular and spectral scattering in the Rayleigh-Gans-Born approximation for particles of various geometric shapes. *Appl. Opt.*, 18:126–130, 1979.

[131] N. Wolfson, J. H. Joseph, and Y. Mekler. Comparative study of inversion techniques. Part I: Accuracy and stability. *J. Appl. Meteorology*, 18:543–555, 1979.

[132] N. Wolfson, Y. Mekler, and J. H. Joseph. Comparative study of inversion techniques. Part II: Resolving power, conservation of normalization and superposition principles. *J. Appl. Meteorology*, 18:556–561, 1979.

[133] G. Yamamoto and M. Tanaka. Determination of aerosol size distribution from spectral attenuation measurements. *Appl. Opt.*, 8:447–453, 1969.

[134] D. L. Jaggard and Y. Kim. Diffraction by band-limited fractal screens. *J. Opt. Soc. Am. A*, 4:1055–1062, June, 1987.